



**UNIVERSIDAD
DE MÁLAGA**

Facultad de Ciencias

Departamento de Biología Molecular y Bioquímica

Programa del doctorado: Biología Celular y Molecular



TESIS DOCTORAL

**Análisis funcional de las asparraginasas
de pino (*Pinus pinaster* Aiton)**

Autor: Sonia Van Kerckhoven


**Directores: Francisco M. Cánovas Ramos
Francisco J. Ruiz Cantón**

2017



UNIVERSIDAD
DE MÁLAGA

AUTOR: Sonia Helena Van Kerckhoven

 <http://orcid.org/0000-0003-2775-8391>

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional:

<http://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

Cualquier parte de esta obra se puede reproducir sin autorización
pero con el reconocimiento y atribución de los autores.

No se puede hacer uso comercial de la obra y no se puede alterar, transformar o hacer obras derivadas.

Esta Tesis Doctoral está depositada en el Repositorio Institucional de la Universidad de Málaga (RIUMA): riuma.uma.es

Dr. FRANCISCO M. CÁNOVAS RAMOS Y Dr. FRANCISCO J. RUIZ CANTÓN,
Profesores del Departamento de Biología Molecular y Bioquímica de la
Universidad de Málaga

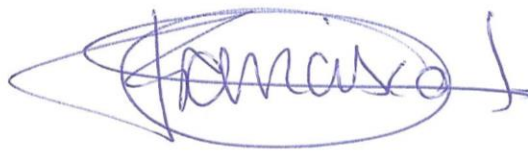
Certifican:

Que la Tesis Doctoral titulada "Análisis funcional de las asparraginasas de pino (*Pinus pinaster* Aiton)", presentada por Sonia Van Kerckhoven para optar al grado de Doctor por la Universidad de Málaga, ha sido realizada bajo nuestra dirección en el Departamento de Biología Molecular y Bioquímica de la Universidad de Málaga y reúne los requisitos exigidos para su defensa ante el tribunal.

En Málaga a 26 de junio de 2017



Francisco M. Cánovas Ramos



Francisco J. Ruiz Cantón

Memoria de tesis doctoral presentada por

Sonia Van Kerckhoven

Para optar al grado de Doctor en Biología por la Universidad de Málaga

Directores:

Dr. Francisco M. Cánovas Ramos y Dr. Francisco J. Ruiz Cantón

Lugar:

Departamento de Biología Molecular y Bioquímica de la Universidad de Málaga

Proyectos que la han financiado:

AP2010-5434 del Ministerio de Educación

BIO2015-69285-R del Ministerio de Economía y Competitividad

BIO-474 de la Junta de Andalucía



Francisco M. Cánovas Ramos



Francisco J. Ruiz Cantón

Málaga, 26 de junio de 2017

A mis padres

ÍNDICE

ABREVIATURAS	11
1. INTRODUCCIÓN.....	15
1.1 Investigación en asparraginasas.....	15
1.2 Filogenia de las asparraginasas de plantas.....	17
1.3 Estructura y activación de las asparraginasas de plantas	21
1.3.1 Características de las enzimas de la familia NTN hidrolasas	21
1.3.2 Activación de las ASPG de plantas.....	23
1.4 Actividad asparraginasa / isoaspartilo dipeptidasa	26
1.4.1 Mecanismo de hidrólisis de la ASPG de plantas	26
1.4.2 Efecto del ion potasio sobre la actividad.....	27
1.4.3 Función de la región variable al extremo carboxilo de la subunidad α	28
1.5 Función de las asparraginasas en plantas	29
1.5.1 Papel de los aminoácidos de la familia del aspartato.....	29
1.5.2 Función de las asparraginasas en especies angiospermas.....	32
1.5.3 Metabolismo de nitrógeno y función de la ASPG en gimnospermas .	33
1.6 <i>Pinus pinaster</i> Aiton como especie modelo.....	37
1.6.1 Filogenia, distribución geográfica y morfología del árbol.....	37
1.6.2 Valor ecológico y económico de coníferas.....	38
1.6.3 Interés de la especie <i>Pinus pinaster</i> Aiton como modelo.....	39
2. ANTECEDENTES Y OBJETIVOS	41
2.1 Antecedentes.....	41
2.2 Objetivos.....	42
3 MATERIAL Y MÉTODOS	43

3.1 Material biológico	43
3.2 Condiciones experimentales de crecimiento de las plantas	46
3.3 Extracción de RNA y síntesis de cDNA	47
3.4 Clonación de las secuencias completas de cDNA de los genes de asparraginasas en pino	49
3.5 Construcción de vectores para la expresión de la proteína recombinante en hojas de <i>N. benthamiana</i>	52
3.6 Expresión <i>in planta</i> y extracción de proteínas recombinantes	55
3.7 Purificación de proteínas recombinantes de pino, expresadas en <i>N. benthamiana</i>	57
3.8 Construcción del vector para la expresión de la proteína recombinante en <i>E. coli</i>	58
3.9 Inducción de la expresión y purificación de la proteína recombinante de <i>E. coli</i>	59
3.10 Análisis por SDS-PAGE y western blot	61
3.11 Ensayo de actividad asparraginasas/isoaspartil dipeptidasa	62
3.12 Caracterización de las proteínas recombinantes purificadas	63
3.13 Espectrometría MALDI-TOF	65
3.14 Extracción de DNA y aislamiento de la región promotora <i>PpASPG1</i>	65
3.15 Análisis <i>in silico</i> de la región promotora <i>PpASPG1</i>	67
3.16 Análisis funcional de la región promotora <i>PpASPG1</i>	68
3.17 Aislamiento de la región promotora <i>PpASPG1</i> en pinos de diferentes procedencias	69
3.18 Reacciones de PCR cuantitativo en tiempo real	70
4. RESULTADOS	75
4.1 Identificación de asparraginasas en el transcriptoma de <i>Pinus pinaster</i> Aiton.	75

4.1.1 Aislamiento de secuencias de cDNA de longitud completa	75
4.1.2 Estructura primaria de las ASPG de coníferas	76
4.1.3 Clasificación filogenética de las asparraginasas de pino.....	80
4.2 Actividad catalítica de las tres asparraginasas de pino	81
4.2.1 Expresión de las proteínas recombinantes <i>in planta</i> y comparación del proceso de maduración	81
4.2.2 Determinación de la preferencia de sustrato.....	84
4.3 Estructura de la PpASPG1 madura y mecanismo de activación	88
4.3.1 Análisis de la proteína PpASPG1 madura	88
4.3.2 Efecto de alteraciones en la RV sobre el proceso de maduración	92
4.3.3 Capacidad de autoprocesamiento de PpASPG1.....	95
4.4 Efecto de la RV sobre las características bioquímica de PpASPG1	99
4.4.1 Efecto de la supresión de la RV sobre las características catalíticas de PpASPG1.....	99
4.4.2 Contribución de la RV a la estabilidad de la enzima PpASPG1.....	103
4.4.3 Papel del ion potasio en la actividad de PpASPG1 y relación con la RV	104
4.5 Caracterización de la región promotora de <i>PpASPG1</i>	107
4.5.1 Aislamiento de la región promotora del gen <i>PpASPG1</i>	107
4.5.2 Análisis <i>in silico</i> de la secuencia promotora	110
4.5.3 Análisis funcional de la región promotora	112
4.5.4 Variabilidad genómica en los promotores de <i>PpASPG1</i> de diferentes poblaciones de <i>P. pinaster</i>	114
4.6 Perfiles de expresión de los genes de asparraginasas de pino	117
4.6.1 Perfiles transcripcionales de los tres genes de asparraginasas durante las fases de desarrollo temprano del pino.....	117

4.6.2. Respuesta de los genes <i>ASPG</i> , δ - <i>OAT</i> y <i>GS1b</i> a la disponibilidad de nitrógeno	124
4.6.3 Efecto de factores de estrés sobre la expresión de <i>PpASPG1</i> y <i>PpASPG2</i> en plántulas.....	128
5. DISCUSIÓN	133
5.1 Identificación y caracterización de asparraginasas de <i>Pinus pinaster</i> ..	133
5.1.1 El transcriptoma de pino marítimo revela la presencia de tres genes para ASPG.....	133
5.1.2 Las tres ASPG de pino presentan doble actividad asparraginasasa/isoaspartilo dipeptidasa	134
5.2 Características particulares de la proteína PpASPG1 y función de la región variable	138
5.2.1 Las asparraginasas de coníferas del tipo dependiente de K ⁺ presentan la región variable de mayor extensión entre las ASPG caracterizadas	138
5.2.2 Modificaciones en la región variable afectan a la maduración del precursor <i>in planta</i> e <i>in vitro</i>	139
5.2.3 Papel de la región variable en las características bioquímicas de PpASPG1.....	141
5.2.4 Efecto del ion potasio sobre la estabilidad y la actividad catalítica de PpASPG1.....	143
5.3 Características de la región promotora de <i>PpASPG1</i>	145
5.4. Análisis de los niveles de mRNA para asparraginasas en el pino	148
5.4.1 Función de las ASPG durante el desarrollo de pino	148
5.4.2 Implicación de la asparraginasasa en la adaptación del metabolismo de pino a cambios ambientales.....	151
6. CONCLUSIONES.....	157
7. BIBLIOGRAFÍA.....	159

8. ANEXOS.....	171
Suplemento 1: Referencias de las bases de datos de secuencias primarias de ASPG usadas en el análisis filogenético	171
Suplemento 2: Cebadores usados en las reacciones de PCR	173
Suplemento 3: Contribución que avala la tesis doctoral	175
Suplemento 4: Otras contribuciones.....	191

ABREVIATURAS

AAT	Aspartato aminotransferasa
ABA	Ácido abscísico
ALL	Leucemia linfoblástica aguda
Arg	Arginina
AS	Asparragina sintetasa
Asp	Aspartato
Asn	Asparragina
ASPG	Asparraginasa
ATP	Adenosina trifosfato
BLAST	<i>Basic Local Alignment Search Tool</i>
BSA	Albúmina de suero bovino
cDNA	DNA complementario
Ct	Ciclo umbral
CTAB	Bromuro de cetil trimetil amonio
Cys	Cisteína
DNA	Ácido desoxirribonucleico.
dNTP	Deoxi nucleótido trifosfato
EcAIII	<i>Escherichia coli</i> producto del gen iaaA
EDTA	Ácido etilén-diamino-tetraacético
EST	Marcador de secuencia expresada (<i>Expressed sequence tag</i>)

Fd-GOGAT	Glutamato sintasa dependiente de ferredoxina
FL	Secuencia completa (<i>Full length</i>)
GDC/SHMT	Complejo glicina descarboxilasa/serina hidroximetiltransferasa
Glu	Glutamato
Gln	Glutamina
Gly	Glicina
GOT	Glutamato oxalacetato transaminasa
GS	Glutamina sintetasa
HEPES	Ácido N-2-hidroxietilpiperazina -N-2-etanosulfónico.
His	Histidina
ICDH	Isocitrato deshidrogenasa
IPTG	Isopropil- β -D-tiogalactopiranósido
JA	Ácido jasmónico
LB	Medio de cultivo Luria-Bertani
LIA	Asparraginasa de <i>Lupinus luteus</i>
Lys	Lisina
MDH	Malato deshidrogenasa
MeJA	Metil jasmonato
MES	Ácido 2-(N-morfolino) etanosulfónico
Met	Metionina
mRNA	RNA mensajero
MOPS	Ácido 3-(N-morfolino) propanosulfónico

N	Nucleótido inespecífico
NaAc	Acetato de sodio
NADH	Nicotinamida adenina dinucleótido reducido
nt	Nucleótido
NTN	N-terminal nucleófilo
OA	Oxalacetato
δ-OAT	Ornitina-δ-aminotransferasa
2-OG	2-oxoglutarato
ORF	Marco abierto de lectura
PAGE	Electroforesis en gel de poliacrilamida
PAL	Fenilalanina amonio liasa
PAT	Prefenato aminotransferasa
pb	Pares de bases
PCR	Reacción de la polimerasa en cadena
Phe	Fenilalanina
PMSF	Fluoruro de fenilmetilsulfonato
Poli-A	Poli-adenosinas
PVP	Polivinilpirrolidona
qRT-PCR	PCR cuantitativa en tiempo real
RLM-RACE	Amplificación rápida de los extremos de cDNA mediada por RNA ligasa
RNA	Ácido ribonucleico

RV	Región con una secuencia variable al extremo carboxilo de la subunidad α en la estructura de las asparraginasas de plantas
SAM	S-adenosil metionina
SAMS	SAM Sintetasa
SD	Desviación estándar
SDS	Dodecil sulfato sódico
Ser	Serina
SNP	Polimorfismo de nucleótido único (<i>Single Nucleotide Polymorphisms</i>)
TAIR	<i>The Arabidopsis Information Resource</i>
TBS	Tampón Tris salino
TEMED	NNN 'N ' tetrametilén-diamina
Thr	Treonina
TMV	Virus del mosaico de tabaco
UPGMA	Análisis de agrupamientos por ligamiento promedio (<i>Unweighted Pairgroup Method with Arithmetic Averages</i>)
UTR	Parte del cDNA no traducida
X-gal	5-bromo-4-cloro-3-indolil- β -D-galactósido

1. INTRODUCCIÓN

1.1 Investigación en asparraginasas

Las enzimas anotadas como asparraginasas (ASPG, EC 3.5.1.1) catalizan la hidrólisis del grupo amido de la asparagina liberando como productos aspartato y amonio. Algunas enzimas además tienen la capacidad de catalizar sustratos con modificaciones tales como glucosilaciones y dipéptidos de isoaspartilo pero el producto común de todas las reacciones es el aspartato. Lang (1904) detectó la enzima por primera vez en tejidos bovinos y en 1922 Clementi confirmó la presencia de actividad asparraginasa en casi todos los tejidos de herbívoros y en hígado de omnívoros. Kidd (1953) descubrió que el suero de cobaya tiene una propiedad antineoplásica la cual más tarde fue atribuida a la actividad asparraginasa por Broome (1963). La EcAII de *E. coli*, una ASPG que es secretada por la bacteria, fue la primera asparraginasa usada clínicamente en un tratamiento contra la leucemia linfoblástica aguda (ALL) (Roberts, 1976). Asparagina es un aminoácido esencial para las células malignas porque carecen de actividad asparagina sintetasa (AS, EC 6.3.5.4) y por ende dependen de una aportación externa. La ASPG ayuda a eliminar la asparagina de la sangre y de tal manera privar a las células cancerígenas de esta molécula imprescindible para su crecimiento (Ho *et al.*, 1970). En humanos se expresa una asparraginasa (hASNase3) que tiene una K_m para asparagina de 3,4 mM (Cantor *et al.*, 2009). La concentración de asparagina en el suero es aproximadamente 50 mM (Nomme *et al.*, 2012) y para que la terapia sea eficiente se necesita una enzima con mayor afinidad para el sustrato, en el orden de micromolar, como las ASPG que se encuentran en las bacterias. La terapia con las enzimas de origen bacteriano produce intolerancias y efectos secundarios no deseados en muchos pacientes con ALL, por lo cual la mayoría de la investigación sobre estas enzimas se centra en la búsqueda de otros organismos fuentes de ASPG como plantas y hongos y/o la mejora de sus características catalíticas.

Otra asparraginasa producida por las bacterias como la EcAI de *E. coli* es de localización citosólica y no presenta propiedades antineoplásicas debido a su baja afinidad para asparragina (Willis and Woolfolk, 1974). En 2000, Borek y Jaskolski publicaron la purificación de un tercer tipo de ASPG de células de *E. coli* denominada EcAIII que no está relacionada evolutivamente con las ASPG de tipo I y II bacterianas. Esta enzima pertenece a la familia de NTN hidrolasas y tiene alta similitud con la hASNase3 y las asparraginasas que se encuentran en las plantas. Las primeras plantas en las que se detectó actividad ASPG fueron leguminosas (Lees and Blakeney, 1970), y hasta hoy en día, la mayoría de asparraginasas de plantas estudiadas pertenecen a este grupo taxonómico. La tabla 1.1 contiene una lista con los principales trabajos que han sido publicados sobre las asparraginasas de plantas.

Estudios realizados en los años 1970 – 1980 detectaron actividad asparraginasa en semillas, hojas jóvenes y nódulos en especies de la familia *fabaceae*. Sodek *et al.* (1980) fue el primero en purificar la enzima de diferentes especies y tejidos vegetales y observó la existencia de dos subgrupos dentro de las ASPG de plantas, dependiente o independiente del ion potasio para su actividad. El nivel de expresión de la isoforma dependiente de potasio es generalmente inferior al nivel de expresión de su parólogo y los transcritos de ambas formas se detectan sólo en determinados tejidos (Ivanov *et al.*, 2012). Debido al bajo nivel de estas proteínas en tejidos vegetales, especialmente en el caso de la ASPG dependiente de potasio (Chang and Farnden, 1981), la mayoría de los trabajos publicados son de proteínas recombinantes expresadas en *E. coli* y tratan de especies leguminosas. En 2006, Bruneau *et al.* confirmaron que *Arabidopsis thaliana* tiene dos asparraginasas que pertenecen a subfamilias de NTN hidrolasas evolutivamente distintas. Los autores que han caracterizado los dos subtipos de una misma especie detectaron una doble actividad asparraginasa/isoaspartilo dipeptidasa en las ASPG independientes de potasio mientras que el subtipo dependiente del ion sólo pudo hidrolizar asparragina (Bruneau *et al.*, 2006; Credali *et al.*, 2011). No obstante Gabriel *et al.* (2012) registraron también una doble actividad para la ASPG potasio dependiente de

A. thaliana. Los estudios de cristalización y la publicación de sus estructuras han arrojado más luz sobre los procesos de activación y catálisis de las ASPG de plantas. Sin embargo, las asparraginasas de especies gimnospermas apenas aparecen en las referencia bibliográficas, con la excepción de unos pocos trabajos (Cañas *et al.*, 2007; Canales *et al.*, 2010; Flores-Monterroso *et al.*, 2013). En cualquier caso, no hay trabajos publicados en los que se realice una caracterización en detalle de la proteína de gimnospermas.

1.2 Filogenia de las asparraginasas de plantas

Las asparraginasas de plantas como miembros de la familia de NTN hidrolasas son estructuralmente diferentes a las ASPG bacterianas y con un origen evolutivo diferente (Michalska and Jaskólski, 2006). La figura 1.1 representa un árbol filogenético de las ASPG de plantas de varias especies cuyas secuencias están disponibles en las bases de datos e incluye la EcAIII de *E. coli* y la hASNase3 humana que pertenecen a la misma familia. En los genomas de briofitas existen varias secuencias codificantes para ASPG cuyas secuencias primarias tienen un porcentaje de identidad media de 83 % entre ellas y 57% con sus homólogas en plantas vasculares. Las bases de datos con EST de plantas terrestres primitivas tales como *Ceratopteris richardii* y las gimnospermas *Ginkgo biloba* y *Cycas rumphii* no contienen secuencias con una similitud significativa, probablemente debido al bajo nivel de expresión de las ASPG en general. Por otro lado, la secuenciación de los genomas de coníferas es tediosa debido a la gran cantidad de pseudogenes, secuencias repetitivas y elementos móviles diversificados (Mackay *et al.*, 2012) y las secuencias genómicas disponibles para ASPG en estas especies resultan mayoritariamente incompletas.

Tabla 1.1 Principales trabajos publicados sobre asparraginasas de plantas.

(?) no especifica si se trata del subtipo dependiente o independiente de potasio.

Especie	Enzima	Referencia
Estudios <i>in planta</i>		
<i>Lupinus albus</i>	ASPG (?)	Atkins <i>et al.</i> , 1975
<i>Glycine max</i>	ASPG (?)	Streeter, 1977
<i>Pisum spp.</i>	ASPG (ambos subtipos)	Sodek <i>et al.</i> , 1980
<i>Lupinus spp</i>		
<i>Vicia faba</i>		
<i>Phaseolus multiflorus</i>		
<i>Zea mays</i>		
<i>Hordeum vulgare</i>		
<i>Lupinus arboreus y angustifolius</i>	ASPG (ambos subtipos)	Chang and Farnden, 1981
<i>Pisum sativum</i>	ASPG (?)	Ireland and Joy, 1983
<i>Lupinus spp.</i>	ASPG (ambos subtipos)	Lea <i>et al.</i> , 1984
<i>Pisum sativum</i>	ASPG (K ⁺ dependiente)	Sieciechowicz <i>et al.</i> , 1985
<i>Lupinus arboreus</i>	ASPG (ambos subtipos)	Lough <i>et al.</i> , 1992
<i>Lupinus polyphyllus</i>	ASPG (K ⁺ independiente)	Sodek and Lea, 1993
<i>Pisum sativum</i>	ASPG (K ⁺ dependiente)	
<i>Lupinus angustifolius</i>	ASPG (?)	Grant and Bevan, 1994
<i>Pisum sativum</i>	ASPG (K ⁺ dependiente)	Chagas and Sodek, 2001
<i>Vigna unguiculata</i>	ASPG (?)	Ali, 2009
<i>Withania somnifera</i>	ASPG (?)	Oza <i>et al.</i> , 2009
<i>Pisum sativum</i>	ASPG (K ⁺ dependiente)	Khalaf <i>et al.</i> , 2012
Expresión en células de <i>E. coli</i>		
<i>Lupinus angustifolius</i>	ASPG (?)	Dickson <i>et al.</i> , 1992
<i>Arabidopsis thaliana</i>	ASPG (K ⁺ independiente)	Casado <i>et al.</i> , 1995
<i>E. coli</i>	EcAIII	Hejazi <i>et al.</i> , 2002
<i>Arabidopsis thaliana</i>	ASPG (K ⁺ independiente)	
<i>E. coli</i>	EcAIII	Borek <i>et al.</i> , 2004
<i>Lupinus luteus</i>	ASPG (K ⁺ independiente)	
<i>Arabidopsis thaliana</i>	ASPG (ambos subtipos)	Bruneau <i>et al.</i> , 2006
<i>Pinus sylvestris</i>	ASPG (K ⁺ dependiente)	Cañas <i>et al.</i> , 2007
<i>Glycine max</i>	ASPG (K ⁺ dependiente)	Cho <i>et al.</i> , 2007
<i>Homo sapiens</i>	hASRGL1	Cantor <i>et al.</i> , 2009
<i>Lotus japonicus</i>	ASPG (ambos subtipos)	Credali <i>et al.</i> , 2011
<i>Arabidopsis thaliana</i>	ASPG (ambos subtipos)	Gabriel <i>et al.</i> , 2011
<i>Withania somnifera</i>	ASPG (?)	Oza <i>et al.</i> , 2011
<i>Homo sapiens</i>	hASRGL1 / hASNase3	Karamitros and Konrad, 2014
Cristalización		
<i>Lupinus luteus</i>	ASPG (K ⁺ independiente)	Michalska <i>et al.</i> , 2006
<i>E. coli</i>	EcAIII	Michalska <i>et al.</i> , 2008
<i>Homo sapiens</i>	hASRGL1 / hASNase3	Nomme <i>et al.</i> , 2012
<i>Homo sapiens</i>	hASRGL1 / hASNase3	Su <i>et al.</i> , 2013
<i>Phaseolus vulgaris</i>	ASPG (K ⁺ dependiente)	Bejger <i>et al.</i> , 2014
<i>Homo sapiens</i>	hASRGL1 / hASNase3	Nomme <i>et al.</i> , 2014

En el genoma de *Picea abies* por ejemplo hay un total de tres genes con una alta homología con las dos asparraginasas de *Arabidopsis thaliana*, pero ninguna secuencia se ha incluido en el análisis filogenético debido a la falta de alguna(s) parte(s) codificante(s). En el genoma de *Pinus taeda* sólo uno de un total de tres genes está completamente secuenciado y en el árbol filogenético se agrupa con una ASPG de *Pinus sylvestris*. En la figura 1.1 se aprecia claramente la separación de los dos subgrupos de asparraginasas dependientes e independientes de potasio. Todas las angiospermas tienen, al menos, una enzima de cada subtipo y la poliploidía debido a eventos de duplicación genómica se refleja en un mayor número de ASPG para ciertas especies. La diferenciación de los dos subgrupos de ASPG de plantas, con una isoforma presentando un mecanismo de activación por potasio, probablemente conlleva una especialización de función y actividad como ya ha sido determinado experimentalmente (Sodek *et al.*, 1980; Bruneau *et al.*, 2006; Credali *et al.*, 2011; Gabriel *et al.*, 2012). Debido a la falta de secuencias disponibles de plantas terrestres primitivas es difícil identificar el momento evolutivo de la separación entre los subgrupos. La poca información disponible sobre las ASPG en gimnospermas no deja claro si dentro de este grupo también existen los dos subtipos de enzimas. Dentro de las plantas leñosas la composición y estructura del sistema vascular de gimnospermas difiere de angiospermas y las asparraginasas podrían tener diferentes papeles entre las diferentes especies de plantas.

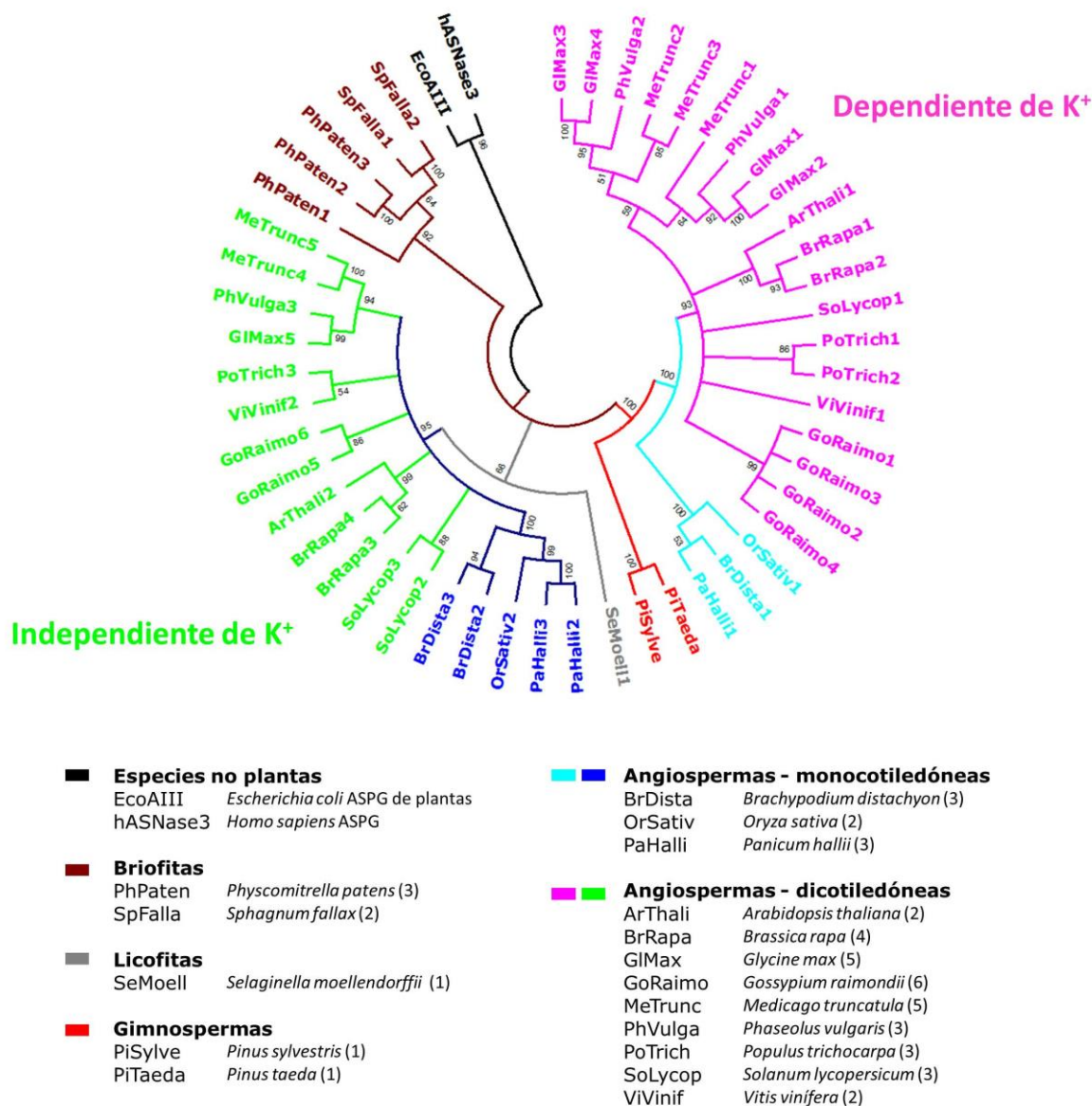


Figura 1.1 Árbol filogenético de asparagininasas de plantas. Se buscaron secuencias codificantes para ASPG de plantas en especies cuyo genoma ha sido secuenciado y las secuencias traducidas se alinearon mediante la aplicación T-Coffee (Notredame *et al.*, 2000). Las ASPG EcoAIII de *E. coli* y hASNase3 de *H. sapiens* se incluyeron como grupo externo. El árbol filogenético se elaboró con el programa MEGA versión 6 (Tamura *et al.*, 2013) por el método UPGMA y representa el resultado de 1000 repeticiones. La leyenda muestra las especies incluidas y el número de ASPG que están representados en el árbol entre paréntesis. Las referencias correspondientes de entradas en los bancos de datos se encuentran en el suplemento 1.

1.3 Estructura y activación de las asparraginasas de plantas

1.3.1 Características de las enzimas de la familia NTN hidrolasas

Los primeros estudios cristalográficos de asparraginasas se llevaron a cabo sobre la asparraginasa bacteriana EcAII (Epp *et al.*, 1971), pero la estructura completa no se publicó hasta 1993 por Swain *et al.* Una vez disponible el modelo de EcAII múltiples estructuras de ASPG bacterianas han sido depositadas en las bases de datos, pero desgraciadamente las ASPG de tipo eucariota que pertenecen a la familia de NTN hidrolasas no han gozado del mismo interés. Las NTN hidrolasas representan una superfamilia de amidohidrolasas que incluye enzimas con actividades diversas tales como asparragina sintetasa, glucosamina 6 fosfato sintasa, N-acilhomoserina lactona acilasa, IMP ciclohidrolasa (Galperin and Koonin, 2012) y existen en diferentes formas oligoméricas como por ejemplo un simple heterodímero en el caso de la penicilina acilasa hasta las 14 subunidades del proteosoma (Michalska and Jaskólski, 2006). Fueron caracterizadas por Brannigan *et al.* (1995) quien propuso el nombre de NTN hidrolasas porque comparten un plegamiento típico con un nucleófilo al extremo amino de una lámina β (N-terminal nucleophile) posicionado de tal manera que permite un ataque nucleofílico además de un procesamiento autocatalítico. La estructura típica consiste en un núcleo catalítico de cuatro capas en un plegamiento tipo $\alpha\beta\beta\alpha$ donde dos láminas β antiparalelas están enfrentadas y flanqueadas a ambos lados por capas de hélices α (Oinonen and Rouvinen, 2000). Las enzimas de la familia NTN hidrolasas son sintetizadas como precursores inactivos y el procesamiento autocatalítico consiste en la eliminación del primer aminoácido, de un propéptido o bien en la separación del precursor en subunidades por una endoproteólisis que puede o no seguir con la eliminación de un péptido (Michalska and Jaskólski, 2006). Son enzimas evolutivamente relacionadas, pero con un grado de diversificación tal que no comparten similitud reconocible en sus secuencias. La conservación de secuencia se limita a un residuo

catalítico serina, cisteína o treonina que queda posicionado en el N-terminal de la subunidad β después del proceso de la maduración.

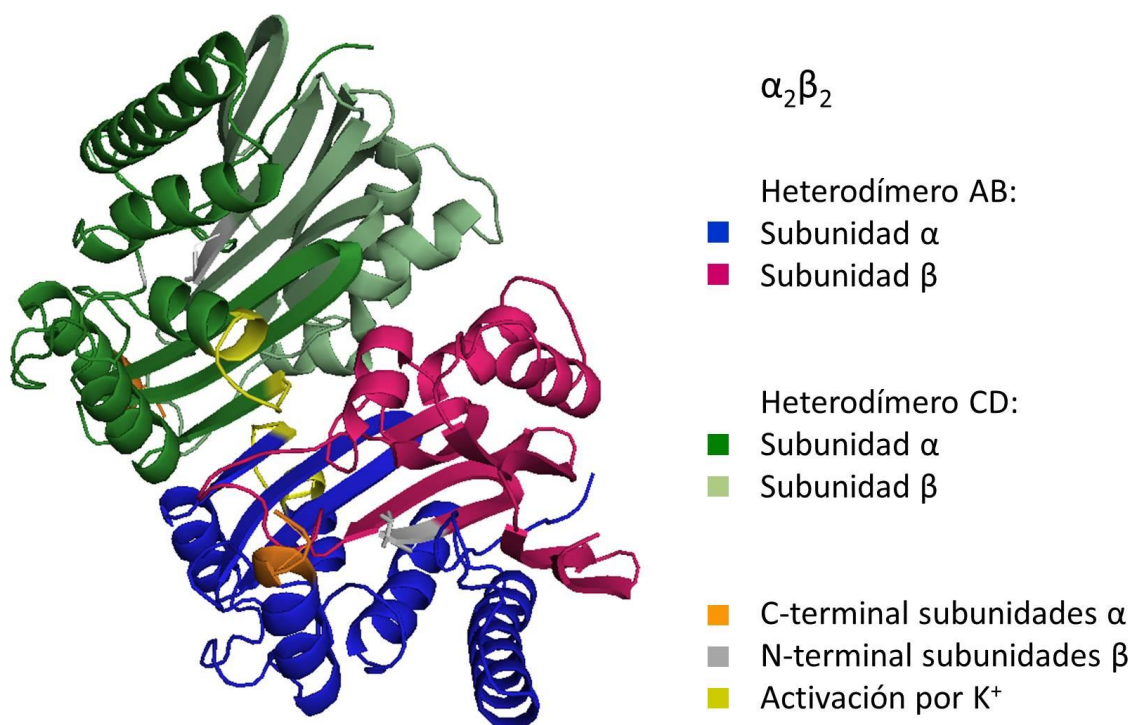


Figura 1.2 Modelo de la estructura heterotetramérica de la asparraginasa de plantas. El modelo se construyó mediante la aplicación SWISS-MODEL (Biasini *et al.*, 2014) sobre la estructura de la ASPG dependiente de potasio de *Phaseolus vulgaris* disponible en la base de datos (Bejger *et al.*, 2014) y se representó con software "The PyMOL Molecular Graphics System, Version 1.8" (Schrödinger, LLC.). Las subunidades α y β que resultan del procesamiento del primer precursor (heterodímero AB) están marcadas en azul y rojo respectivamente. La región que interactúa con el ion potasio en color amarillo. Las subunidades del segundo heterodímero CD están en verde. La secuencia variable al extremo carboxilo de las subunidades α no se incluyó en el modelo. El C-terminal de la secuencia conservada de subunidades α está marcada en color naranja para marcar la posición dónde empezaría la región variable. El extremo N-terminal de las subunidades β que empieza con el residuo nucleófilo (en color gris) marca la posición del sitio catalítico.

En el caso de las L-asparaginasa, aspartilglucosaminidasa y taspasa1 la enzima madura forma un heterotetrámero tipo $(\alpha\beta)_2$. La figura 1.2 representa la estructura de la ASPG dependiente de potasio de *Phaseolus vulgaris* como ejemplo. El procesamiento autocatalítico del precursor genera dos subunidades con el cual el residuo catalítico (treonina) se expone al extremo N-terminal de la subunidad β . La enzima madura consiste en cuatro subunidades que derivan de dos precursores.

1.3.2 Activación de las ASPG de plantas


El mecanismo de activación no es como ocurre con las proenzimas (zimógenos) donde una región de la proteína que inhibe la actividad es eliminada mediante ruptura proteolítica como por ejemplo en el caso de tripsinógeno, pepsinógeno y caspasa. En las NTN hidrolasas el plegamiento de la enzima no es diferente antes y después del procesamiento y las subunidades se mantienen fuertemente unidas entre sí para formar una unidad funcional (Su *et al.*, 2013). El sustrato también puede alojarse en el sitio activo de la ASPG no procesada, pero entonces no se produce la catálisis (Nomme *et al.*, 2012). El autoprocesamiento es una modificación postraducciona en *cis* que requiere de un preciso alineamiento de varios grupos funcionales (Xu *et al.*, 1999). Este mecanismo de maduración se puede comparar con el descrito para serina proteasas (Carter and Wells, 1988) donde un residuo nucleófilo, que en el caso de las ASPG de plantas probablemente es una treonina, forma una triada catalítica juntos con una base (histidina) y un ácido (aspartato). El proceso se inicia con la transferencia del protón del grupo OH de la treonina a la histidina, facilitada por la presencia de un Asp cercano cuya carga negativa estabiliza la protonación de la histidina. El nucleófilo activado ataca el carbonilo del sustrato que genera una transición $N \rightarrow O$ con la formación de un intermediario tetraédrico tipo éster. No obstante, existe controversia sobre qué residuos y grupos funcionales intervienen exactamente en la activación del nucleófilo que promueve la autocatálisis en las ASPG de plantas.

Aparentemente el residuo catalítico que actúa de nucleófilo en la hidrólisis del sustrato (treonina) también interviene al proceso de la autoactivación ya que los precursores con este residuo mutado pierden la capacidad de autoactivarse (Michalska *et al.*, 2008). Cuando se comparan las estructuras primarias de las asparraginasas de diferentes especies (ver la figura 1.3) se observa mucha variabilidad tanto en longitud como en la secuencia de aminoácidos que precede a esta treonina catalítica. Además, los trabajos de cristalización de ASPG de plantas sugieren que el extremo C-terminal de la subunidad α está desestructurada porque tiene muy poca densidad electrónica (Michalska *et al.*, 2006; Bejger *et al.*, 2014). Se ha sugerido que posiblemente los residuos implicados en el mecanismo de autocatálisis difieren entre las diferentes ASPG de plantas (Michalska *et al.*, 2008). La eficiencia del proceso autoproteolítico de las ASPG de plantas también difiere entre las especies y se observa además que las proteínas recombinantes expresadas en sistemas heterólogos como *E. coli* tardan más en activarse (Michalska and Jaskólski, 2006; Cantor *et al.*, 2009). La hASNase3 por ejemplo se procesa lentamente, pero la activación continúa *in vitro* (Nomme *et al.*, 2012). La ASPG dependiente de potasio de *Pinus sylvestris* recuperada de los cuerpos de inclusión de *E. coli* no presentó la capacidad autoproteolítica aunque se observó su activación cuando extractos de radículas de plántulas se añadieron a la proteína recombinante (Cañas *et al.*, 2007). Las asparraginasas de plantas estudiadas pertenecen todas a especies angiospermas, *E. coli* o el ser humano y considerando la incapacidad de la ASPG de *P. sylvestris* de autoactivarse *in vitro*, existe la posibilidad de que en gimnospermas la enzima precisa de una ayuda externa para activarse.

At5g08100 MVGWAIALHGGAGDIPIDLPDERRIPRESALRHCLDLGISALKSGKPPLDVAELVVRELE 60
Psylv -MGWAIHAVHGGAGVDP-CLPMERQQEAKDLLLHCLNLGVSALKASVPAIDVVELLVRELE 58
PvAspG1 MGGWAIHAVHGGAGVDP-NLPPQRQEAKQLLTRCLNLGISSLRNASALDVVELLVRELE 59
At3g16150 MGGWAIHAVHGGAGIDP-NLPAERQEEAKQLLTRCLNLGIIALRSNVSAIDVVELVIRELE 59
 *****:***** * ** :*: :. * :*:**: :*: :*: :*: :*: :*

At5g08100 NHPDFNAGKGSVLTAQGTVEMEASIMDGKTKRCGAVSGLTTVVNPISLARLVMEKTPHIY 120
Psylv TNPVFNSGRGSALSAGKTVEMEASIMDGFRNKCGAVSGLSTVKNPISLARLVMEKSPHIV 118
PvAspG1 TDPLFNSGRGSALTDKGTVEMEASIMDGSNRRCGAVSGLTTVKNPVSLARLVMDKSPHSY 119
At3g16150 TDPLFNSGRGSALTEKGTVEMEASIMDGKTKRCGAVSGITTVKNPISLARLVMDKSPHSY 119
 ..* **:*:**:*: :***** :***** :*: **:***** :*: *

At5g08100 LAFDAAEAFARAHGVETVDSSHFITPENIARLKQAKEFNVRQLDYTVSPKVPD----- 174
Psylv LAFNGAEAFARQGVETVDSYFITAENKERLEQAKSVNSIQFDYRLPESDDAVGKPNEN 178
PvAspG1 LAFNGAEAFARQGVDDVENEYFVTPEENVGMLKLAKEANTILFDYRIPTVGVD----- 172
At3g16150 LAFSGAEDFARKQGVETVDNEYFVTDDNVGMLKLAKEANSILFDYRIPPMGCA----- 172
 ***..** *** :*: :*: :*: :* * :*: * :*: :*

At5g08100 -----NCGD----- 178
Psylv GLLPTVKIHAKGIVAFDKILQNGELSQKELSPLESCMQNGDKMACRKGMLEPFQMNGAPI 238
PvAspG1 -----SCGAGAESPLQMNGLPI 189
At3g16150 -----GA-AATDSPIQMNGLPI 188


At5g08100 --SQIGTVGCVAVDSAGNLASATSTGGYVNKMVGRIGDTPVIGAGTYANHLCAISATGKG 236
Psylv NIYEPETVGCVVVDSEGHCSATSTGGLINKMSGRIGDSPIIAGTYANGLCAVSATGEG 298
PvAspG1 SVYAPETVGCVVVDKEGRCAATSTGGLMNKMSGRIGDSPLIAGTYACGVCVGSCTGEG 249
At3g16150 SIYAPETVGCVVVDGKGHCAAGTSTGGLMNKMMGRIGDSPLIAGTYASEFCVGSCTGEG 248
 ******* * . *:***** :*** ******:***** :*.:*.**:

At5g08100 EDIIRGTVARDVAALMEYKGLSLTEAAAYVVDQSVPRGSCGLVAVSANGEVTPFNTTGM 296
Psylv EAIIRATVGRDVAALMEYKGLSLQESVDFVINKRLEDGKGGLIAVSSNGDVAAGFNTSGM 358
PvAspG1 EAIIRGTLAREVAAMMEYKGLQLQEAVDVFINHRLDEGKAGLIAVSNTGEVAYGFNCNAM 309
At3g16150 EAIIRATLARDVSAVMEYKGLNLQEAVDYVIKHLDEGFAGLIAVSNKGEVVCGFNSNGM 308
 * ***:*:*:*:*:*****.* *: :*: : * ***:** :*.* **..*

At5g08100 FRACASEDGYSEIAIWPNN 315 **independiente de K⁺**
Psylv FRACATEEGYFEVGIWQ-- 375 no caracterizada
PvAspG1 FRGCATEDGFMVVGIWE-- 326
At3g16150 FRGCATEDGFMEVAIWE-- 325 } **dependiente de K⁺**
 .:*:*: :.*

Figura 1.3 Alineamiento de secuencias proteicas de algunas ASPG de plantas. Las secuencias bajadas de las bases de datos se alinearon con el programa Clustal Omega (Sievers *et al.*, 2011). El alineamiento contiene 2 asparraginasas dependientes de potasio (At3g16150 de *A. thaliana*, PvAspG1 de *P. vulgaris*), 1 independiente del ion (At5g08100 de *A. thaliana*) y la secuencia de una ASPG de *Pinus sylvestris* (Psylv) no caracterizada. La región de la subunidad α conservada está marcada con fondo azul, la región variable (RV) en amarillo, el sitio de procesamiento se indica con las tijeras y la subunidad β con fondo de color rosa.

1.4 Actividad asparraginasas / isoaspartilo dipeptidasas

1.4.1 Mecanismo de hidrólisis de la ASPG de plantas

Aunque las ASPG de plantas son estructuralmente similares, sus características enzimáticas son bastante diversificadas (Bejger *et al.*, 2014). Después del estudio de la ASPG independiente de potasio de *Lupinus luteus* y la EcAIII se concluyó que la función primaria de estas enzimas consiste en la actividad isoaspartilo dipeptidasa y que la hidrólisis de asparragina es de importancia secundaria aunque posiblemente beneficiosa para el organismo (Borek *et al.*, 2004). La hASNase3, también tiene mayor preferencia para dipéptidos de isoaspartilo y cataliza asparragina con menor eficiencia (Cantor *et al.*, 2009). Distintos estudios comparativos de ambos subtipos en varias especies de plantas mostraron que las ASPG independientes de potasio presentan la doble actividad mientras que las ASPG dependientes del ion solamente catalizan asparragina (Bruneau *et al.*, 2006; Credali *et al.*, 2011). No obstante, un estudio posterior de la ASPG dependiente de potasio de *Arabidopsis thaliana* determinó que también puede hidrolizar dipéptidos isoaspartilo (Gabriel *et al.*, 2012). La cristalización de la primera ASPG dependiente de potasio de *Phaseolus vulgaris* (Bejger *et al.*, 2014) arroja más luz sobre la especificidad de sustrato y el efecto de los iones sobre la actividad.

El mecanismo de hidrólisis del sustrato por las NTN hidrolasas fue descrito por primera vez por Duggleby *et al.* (1995). A diferencia del mecanismo descrito para las serina proteasas, en vez de una triada catalítica el centro catalítico consiste en un solo aminoácido, que en el caso de las ASPG de plantas es una treonina (Oinonen and Rouvinen, 2000). Este mismo residuo, localizado en el extremo amino de la subunidad β , actúa de nucleófilo y de base catalítica. En la proteína madura el extremo amino de la subunidad β (marcado en color gris en la figura 1.2) se encuentra enterrado en la estructura, entre dos láminas β , y marca el sitio activo en cada heterodímero. El embudo formado por el sitio activo determina la especificidad del sustrato, para asparragina, dipéptidos de

isoaspartilo o aspartilo-glucosamina (Nomme *et al.*, 2012). Bejger *et al.*, (2014) observaron que el bolsillo formado por la cavidad catalítica de PvAspG1 es suficientemente grande para que quepan sustratos con una cadena lateral más grande que asparragina, como por ejemplo dipéptidos de isoaspartilo. La especificidad de sustrato probablemente se determina por la naturaleza de los aminoácidos que se encuentran, en la estructura tridimensional, en el entorno del nucleófilo y de la cadena lateral del sustrato. La figura 1.3 muestra el alineamiento de la secuencia primaria de PvAspG1, las dos ASPG de *A. thaliana*, caracterizadas experimentalmente como dependiente (At3g16150) e independiente (At5g08100) de potasio, y una ASPG de *P. sylvestris* (Psvlv). Se identificaron los 9 residuos que siguen al motivo GAG (Gly11-Ala12-Gly13) en la subunidad α de PvAspG1 como posiblemente relacionados con la especificidad de sustrato y el alineamiento muestra un mayor parecido entre las ASPG dependientes de potasio de angiospermas y Psvlv que con At5g08100.

1.4.2 Efecto del ion potasio sobre la actividad

Las ASPG de ambos subgrupos tienen un bucle de estabilización, localizado en la subunidad α , donde la inserción de un catión metálico (Na^+) provoca la formación de un puente de hidrógeno, en la secuencia PvAspG1 entre la Thr196 y Asn65, importante para la orientación de la treonina durante la catálisis (Bejger *et al.*, 2014). Las ASPG dependientes de potasio tienen además una secuencia que forma un bucle de activación (marcado en color amarillo en la figura 1.2), que también puede alojar un catión tal que la enzima madura de este subtipo coordina un total de 4 cationes metálicos. Según Bejger *et al.*, (2014) el ion que se inserta en el bucle de activación influye sobre el posicionamiento de tres residuos que funcionan como un interruptor catalítico. Las enzimas independientes de potasio tienen la conformación permanente activada mientras que las dependientes se activan por potasio, pero se inactivan cuando tienen sodio unido en el bucle de

activación, aunque mantienen una actividad basal. El potasio es un ion abundante y esencial para la planta cuya concentración es mantenida entre 100 y 200 mM en el citosol para garantizar un metabolismo eficiente. Por otro lado el ion sodio tiene un efecto tóxico para la célula por encima de 10 mM por lo cual normalmente se encuentra a una concentración muy inferior a la del potasio (Kader and Lindberg, 2010). La alta razón K^+/Na^+ mantenida en el citosol sugiere que estos iones no constituyen un mecanismo de control sobre la actividad ASPG, sino que el potasio actúa de manera alostérica e induce un cambio conformacional que aumenta su actividad (Di Cera, 2006). Este cambio conformacional también ayuda a mantener la estabilidad de la enzima (Sodek *et al.*, 1980).

1.4.3 Función de la región variable al extremo carboxilo de la subunidad α

En la comparación de la estructura primaria de las ASPG homólogas mostradas en la figura 1.3 destaca una región de secuencia variable al C-terminal de la subunidad α , que llamamos RV, que en el caso de la ASPG de *P. sylvestris* es más extensa que en las ASPG homólogas de angiospermas (Cañas *et al.*, 2007). La baja densidad electrónica mostrada por la RV en los análisis cristalográficos (Bejger *et al.*, 2014) ha impedido la modelización de su estructura, por lo que se considera que podría ser una región intrínsecamente desestructurada (Michalska *et al.*, 2006). Esto ha dificultado postular una hipótesis sobre su función en base a aspectos estructurales. No obstante, diversas evidencias sugieren que esta región o parte de ella puede tener un papel en el procesamiento autocatalítico, la especificidad de sustrato y/o la activación por el ion potasio. En las asparraginasas de *A. thaliana* eliminaciones y mutaciones en la RV afectan a la autoproteólisis y la actividad catalítica (Gabriel *et al.*, 2012). La ASPG de una arquea termófila presenta una secuencia conectora “linker” covalente que es considerada necesaria para el correcto funcionamiento de proteínas con más que un dominio y para la cual

también se ha definido un papel como chaperona que ayuda al plegamiento y resistencia a la desnaturalización térmica (Tomar *et al.*, 2013). En eucariotas la adición de un conector ha demostrado mejorar la estabilidad de las proteínas y a veces la funcionalidad (Wang *et al.*, 2011). Para procariotas una proteína pequeña presenta una ventaja para la tasa de crecimiento pero con el aumento de complejidad y un menor % individual de proteína con respecto a la masa, el efecto sobre la tasa de crecimiento es mínimo y la longitud no tiene importancia (Wang *et al.*, 2011). La RV extensa que observamos en la ASPG de *P. sylvestris* no tiene homología con ninguna otra secuencia proteica conocida en eucariotas, y se desconoce si puede influir sobre aspectos funcionales de la enzima. Por lo que podría ser de gran interés estudiar si afecta al autoprocesamiento del precursor, si está presente en la enzima madura y, si ese es el caso, tiene algún efecto sobre sus características funcionales.

1.5 Función de las asparraginasas en plantas

1.5.1 Papel de los aminoácidos de la familia del aspartato

La ASPG, junto con la asparragina sintetasa (AS), ocupa una posición clave en la interconversión de asparragina y aspartato. El amonio asimilado por la planta en forma de glutamina es transferido a asparragina por la asparragina sintetasa. Asparragina es el metabolito más abundante en plantas para transporte y almacenamiento de N (Borek and Jaskólski, 2001) y su síntesis compite por el aspartato con las rutas de síntesis de lisina, treonina y metionina (Azevedo *et al.*, 2006). En soja, asparragina constituye el 50% de aminoácidos en cotiledón y la mayor tasa de síntesis de asparragina ocurre ahí (Pandurangan *et al.*, 2012). Es el aminoácido que mayoritariamente es empleado como donador de N para la síntesis de aminoácidos en tejidos sumideros (Gaufichon *et al.*, 2013). El uso de asparragina difiere entre tejidos,

especies y de manera temporal (Lea *et al.*, 2007) y se favorece en órganos en desarrollo, donde tanto el N como el C que provienen de asparragina se destinan a la síntesis de aminoácidos de la familia de Asp (Michalska *et al.*, 2006).

Aparte de su principal papel como constituyentes de las proteínas, los aminoácidos en su forma soluble participan en el metabolismo general de las plantas y actúan de intermediarios en múltiples vías metabólicas. La figura 1.4 resume de manera simplificada las rutas metabólicas relacionadas con los aminoácidos que derivan del aspartato. El aspartato forma parte de la ruta biosintética de purinas y pirimidinas (Stasolla *et al.*, 2003), pero también puede ser catabolizado y el correspondiente esqueleto carbonado entrar en el ciclo del ácido cítrico al igual que todos los aminoácidos de la familia del aspartato. Este ciclo, aparte de generar energía, tiene otras funciones como el suministro de esqueletos de carbono para la síntesis de aminoácidos, isoprenoides, etc. (Kirma *et al.*, 2012). Los aminoácidos treonina, glicina y serina son convertidos fácilmente entre ellos y pueden intervenir tanto en procesos de degradación como de biosíntesis de aminoácidos (Hildebrandt *et al.*, 2015). El catabolismo de glicina en la mitocondria con intervención del complejo glicina descarboxilasa/serina hidroximetiltransferasa (GDC/SHMT) genera serina que es exportada hacia el citosol donde funciona como molécula donadora primaria de grupos metilos para el ciclo de los grupos metilos activados (Cantón *et al.*, 2005). Esta ruta metabólica contribuye al reciclado de SAM, la molécula donadora para reacciones de metilación, como por ejemplo para la síntesis de aminoácidos, poliaminas, fenilpropanoides, etileno y especialmente lignina que es considerado un sumidero muy importante del consumo de SAM en plantas vasculares (Hanson and Roje, 2001). El aspartato además interviene en la ruta de síntesis de arginina, otro compuesto importante para el metabolismo de nitrógeno en plantas (Slocum, 2005).

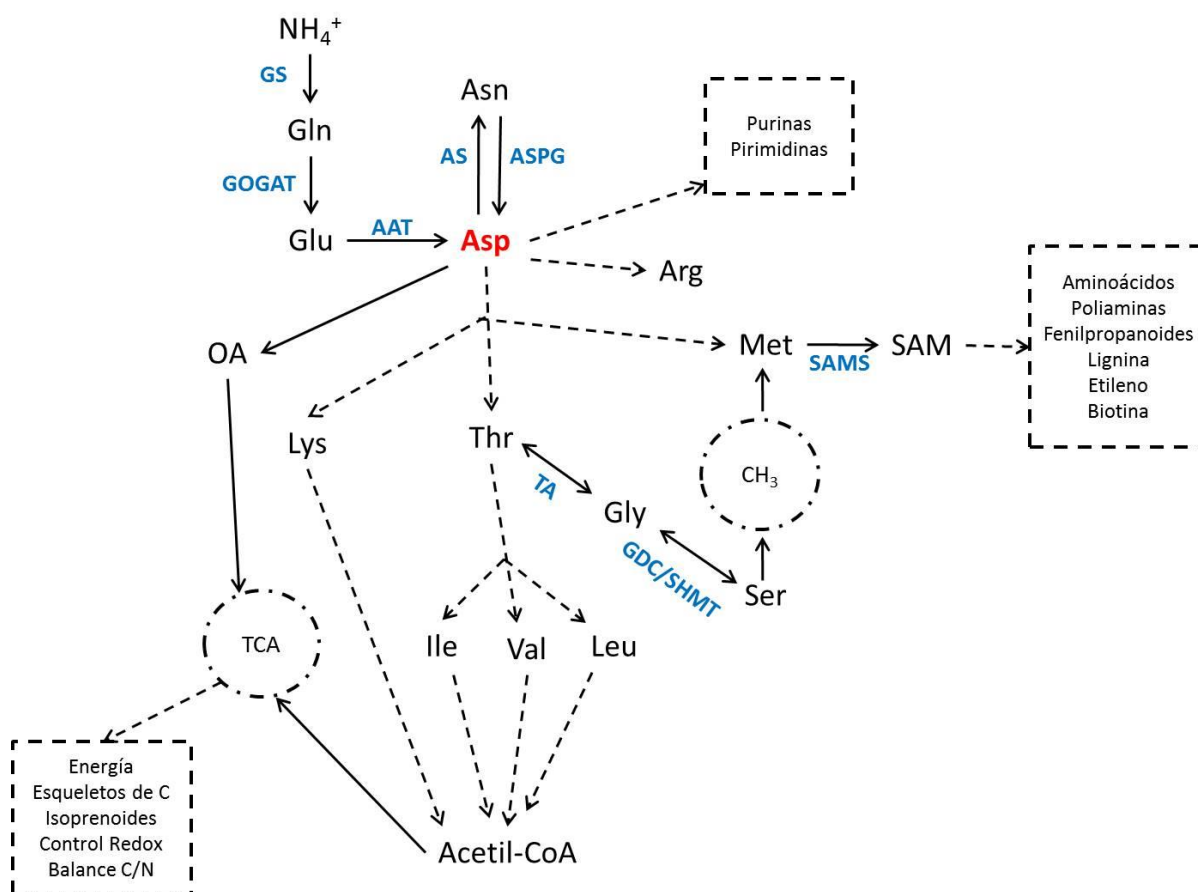


Figura 1.4 Esquema de algunas rutas metabólicas relacionadas con el aspartato. Los aminoácidos y moléculas derivadas están conectados por flechas continuas que representan una reacción de catálisis o flechas discontinuas para rutas metabólicas de varios pasos. Se mencionan algunas enzimas relevantes: GOGAT Glutamato sintasa, AAT aspartato aminotransferasa, AS asparragina sintetasa, ASPG asparraginasa, TA treonina aldolasa, GDC/SHMT complejo glicina descarboxilasa/serina hidroximetil transferasa, SAMS S-adenosilmetionina sintasa. Los círculos representan ciclos metabólicos: TCA ciclo del ácido cítrico, CH_3 ciclo de los grupos metilos activados. Los cuadros mencionan algunos de los destinos y funciones metabólicos de las moléculas o rutas sin especificar con más detalle en este esquema. Figura redibujada y simplificada integrando información de varias publicaciones. (Stasolla *et al.*, 2003; Cantón *et al.*, 2005; Slocum, 2005; Azevedo *et al.*, 2006; Kirma *et al.*, 2012; Hildebrandt *et al.*, 2015).

1.5.2 Función de las asparraginasas en especies angiospermas

Los primeros autores que detectaron actividad asparraginasa en tejidos vegetales (ver tabla 1.1, estudios *in planta*) trabajaron con semillas, hojas jóvenes o nódulos de plantas leguminosas, los tejidos donde más se expresan estas enzimas. La función primaria de las ASPG dependientes de potasio reside en la movilización de N y a las independientes del ion se ha asignado un papel de detoxificación ya que su actividad prioritaria es de isoaspartilo dipeptidasa (Bruneau *et al.*, 2006). La doble actividad asparraginasa/isoaspartilo hidrolasa tiene su sentido en plantas: en las semillas se sintetizan las proteínas de reserva con una gran necesidad para amonio y estas proteínas se tienen que conservar mucho tiempo (Michalska and Jaskólski, 2006). Durante la deshidratación y maduración de la semilla es frecuente que se forman dipéptidos de isoaspartilo espontáneamente que resultan en proteínas no funcionales y agregaciones (Bruneau *et al.*, 2006). Es importante para el desarrollo del embrión que las proteínas incorrectamente sintetizadas sean eliminadas debido al limitado reciclaje y esto requiere una enzima especializada como la asparraginasa.

Cuando una planta llega a ser autótrofa se aumenta la biosíntesis de aminoácidos en las células fotosintéticas. En las hojas los productos de la asimilación de nitrógeno son glutamato, glutamina, aspartato y asparragina por lo que sus concentraciones son muy altas en estos tejidos (Lam *et al.*, 1995; Coruzzi, 2003). Como molécula de transporte la asparragina se dirige hacia tejidos que requieren grandes cantidades de nitrógeno donde este elemento es movilizado por la asparraginasa para la síntesis de proteínas (Sieciechowicz *et al.*, 1988). En general, se detecta mucha actividad ASPG en hojas en desarrollo y su expresión se regula por el ciclo diurno para que el amonio pueda ser eficientemente reasimilado y metabolizado en rutas dependientes de la fotosíntesis (Michalska and Jaskólski, 2006). En hojas jóvenes de guisante el nivel de transcritos de ASPG aumenta en presencia de luz y disminuye en oscuridad (Sieciechowicz *et al.*, 1985). La glutamina

sinetasa (GS, EC 6.3.1.2) también presenta un alto nivel de expresión durante el día y mantiene la concentración de amonio a un nivel reducido por su reasimilación en forma de glutamina (Sieciechowicz *et al.*, 1989). Trabajos de transcriptómica en angiospermas incluyen a la ASPG como un gen cuyos transcritos aumentan en respuesta a una mayor disponibilidad de nitrógeno y posiblemente la razón N/C influye en su nivel de expresión (Scheible *et al.*, 2004; Gutiérrez *et al.*, 2007).

En conclusión, la asparraginasa ocupa una posición clave en el suministro de nitrógeno y aspartato para toda una red de rutas metabólicas. Es especialmente importante en los tejidos sumideros que tienen una alta tasa de biosíntesis, tales como el sistema vascular en las especies leñosas. Es de sospechar que existe mucha diferencia en el control de su expresión entre especies, tejidos y estadios de desarrollo.

1.5.3 Metabolismo de nitrógeno y función de la ASPG en gimnospermas

A lo largo del ciclo de vida de una planta el flujo de aminoácidos cambia considerablemente. El almacenaje de N en semillas y en órganos vegetativos ocurre en forma de proteínas de reserva con un alto contenido en arginina, asparragina y glutamina (Müntz, 1998) y la interconversión de estos aminoácidos juega un papel central en el metabolismo de N en coníferas (Cánovas *et al.*, 2007). Durante la germinación, un proceso que ocurre en oscuridad, las proteínas de reserva de la semilla son degradadas y los aminoácidos liberados sostienen la biosíntesis de las proteínas necesarias para el crecimiento. La energía requerida es obtenida de la oxidación de aminoácidos y la degradación de otros compuestos de reserva como ácidos grasos y almidón hasta que el aparato fotosintético es funcional (Hildebrandt *et al.*, 2015). El flujo de aminoácidos durante este proceso en gimnospermas está bien documentado en la bibliografía y se resume de forma esquemática en la figura 1.5. En semillas de pino las proteínas de reserva contenidas en el

megagametofito son extremadamente ricas en arginina, el aminoácido con el mayor contenido en N (Stone and Gifford, 1997). La movilización de estas proteínas de reserva produce un aumento en el nivel de arginina libre y un flujo hacia el embrión donde el aminoácido es catabolizado a ornitina y urea por la arginasa (EC 3.5.3.1) (Todd *et al.*, 2001). La ornitina es metabolizada por la ornitina- δ -aminotransferasa (δ -OAT, EC 2.6.1.13) que cataliza la transaminación de ornitina a 2-oxoglutarato con la producción de glutamato y glutamato-5-semialdehído (GSA). La urea se hidroliza por la ureasa (EC 3.5.1.5) con la producción de dióxido de carbono y amonio, el cual es reciclado por la vía glutamina sintetasa (GS, EC 6.3.1.2)/glutamato sintasa (GOGAT, EC 1.4.7.1) (Cánovas *et al.*, 1998). Como GOGAT se expresa muy poco en semillas, pero GS se expresa a altos niveles, posiblemente es δ -OAT la enzima que aporta glutamato para soportar el reciclaje activo de amonio. El eficiente reciclaje de amonio liberado durante la germinación por GS contribuye a mantener los niveles de amonio estables en embriones y plántulas (Cañas *et al.*, 2008). La glutamina sintetizada mediante esta ruta metabólica sirve como donadora de N para la biosíntesis de todos los componentes nitrogenados necesarios para el crecimiento de la plántula (Avila *et al.*, 2001). Las coníferas expresan dos isoformas de GS: la GS1a que junto con una glutamato sintasa dependiente de ferredoxina (Fd-GOGAT) se asocia con el desarrollo del cloroplasto en células fotosintéticas y la GS1b de mayor abundancia en hipocótilos y radículas donde al igual que la NADH-GOGAT (EC 1.4.1.14) su expresión se asocia con el sistema vascular (Avila *et al.*, 2001; Lancien *et al.*, 2002; Suárez *et al.*, 2002; Lea and Miflin, 2003; Cánovas *et al.*, 2007). La glutamina generada en las células vasculares de los cotiledones se transporta hacia hipocótilos y radículas donde gran parte del N es canalizado hacia la biosíntesis de asparragina por la AS, el aminoácido más abundante durante el desarrollo de las plántulas de pino (King and Gifford, 1997). El glutamato liberado por la reacción de la AS dependiente de Glutamina es transaminado hacia aspartato por la aspartato amino transferasa (AAT, EC 2.6.1.1). En las plántulas en desarrollo se detectan grandes cantidades de asparragina en hipocótilos (45%) y radículas (19%) debido a una alta actividad AS en estos

tejidos (Cañas *et al.*, 2008). Mientras se van desarrollando las plántulas de pino la expresión de la ASPG se correlaciona espacialmente y temporalmente con la aparición del sistema vascular, lo que sugiere una función para esta enzima en la movilización de N para sostener el desarrollo de este tejido una vez agotadas las reservas del megagametofito (Cañas *et al.*, 2007).

Durante el desarrollo del sistema vascular las coníferas invierten gran cantidad de carbono en la biosíntesis de fenilpropanoides. Estos monómeros constituyen los bloques de construcción para la lignina, un componente de la pared celular que confiere rigor e hidrofobicidad al sistema vascular. Aunque la lignina no contiene N, el esqueleto de fenilpropano de esta ruta metabólica proviene de la deamidación de fenilalanina por la fenilalanina amonio liasa (PAL, EC 4.3.1.5). La ruta del shikimato provee los anillos aromáticos para la biosíntesis de fenilalanina y el grupo amido de este aminoácido es añadido por la prefenato amino transferasa (PAT) por transaminación de glutamato con la liberación de 2-oxoglutarato. No obstante, recientes publicaciones han determinado para esta enzima en *N. benthamiana* también una actividad AAT (de la Torre *et al.*, 2014), pudiendo usar el aspartato como donador del grupo amido con la liberación de oxalacetato. Durante el proceso de la lignificación las células consumen también grandes cantidades de grupos metilos que provienen del metabolismo C1 asociado con el catabolismo de glicina y serina que resulta en la liberación de amonio (Cantón *et al.*, 2005). De tal manera la ASPG podría tener una función importante en proveer el amonio y el aspartato empleado en la biosíntesis de los componentes del sistema vascular.

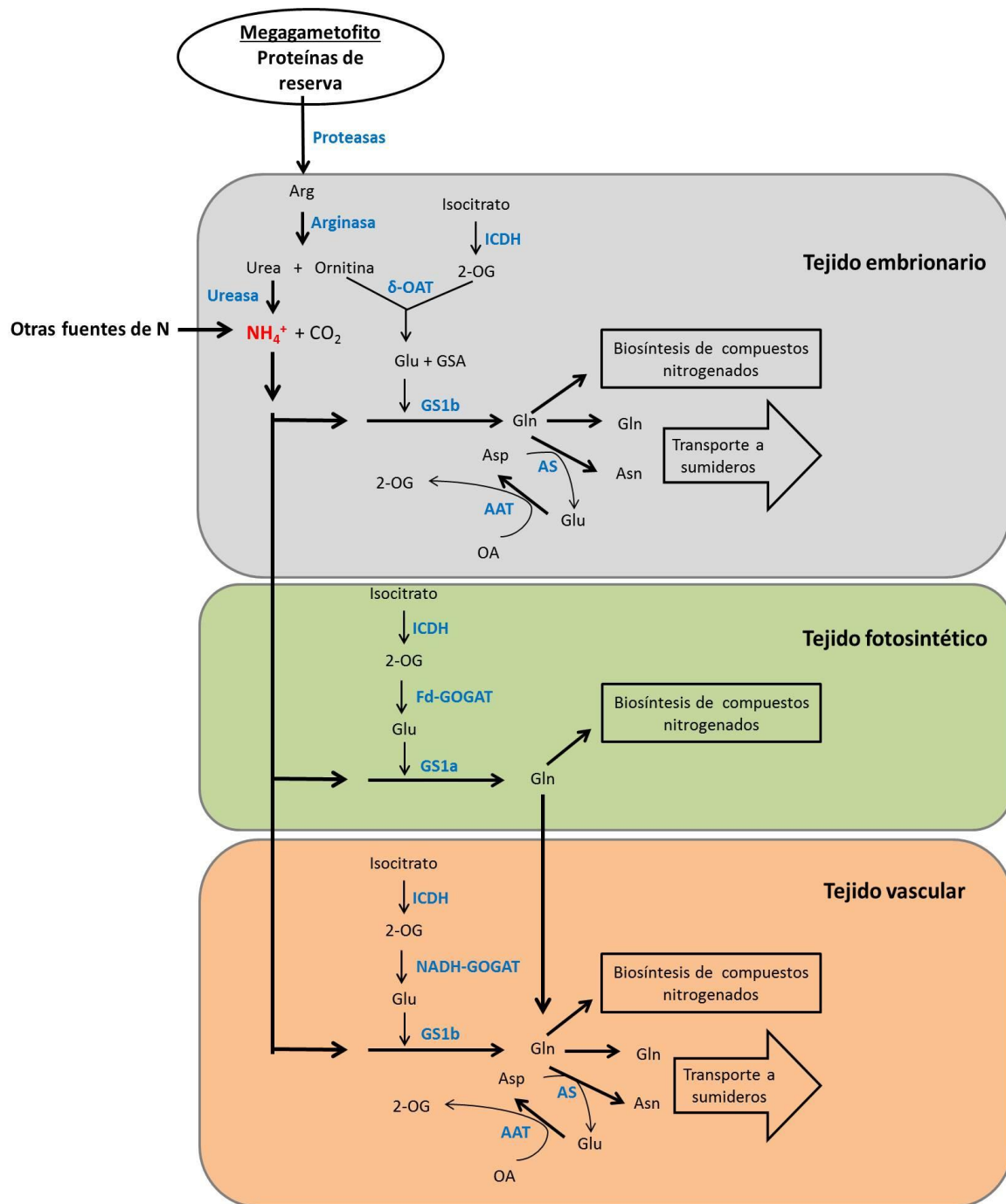


Figura. 1.5 Rutas metabólicas relacionadas con el metabolismo de nitrógeno en coníferas en diferentes tejidos. Nombres de enzimas en azul (ICDH: isocitrato deshidrogenasa, δ -OAT: ornitina- δ -aminotransferasa, GS1b: glutamina sintetasa citosólica b, GS1a: glutamina sintetasa citosólica a, Fd-GOGAT: glutamato sintasa dependiente de ferredoxina, NADH-GOGAT: glutamato sintasa dependiente de NADH, AS: asparagina sintetasa, AAT: aspartato aminotransferasa. Redibujado de Cánovas *et al.*, (2007).

1.6 *Pinus pinaster* Aiton como especie modelo

1.6.1 Filogenia, distribución geográfica y morfología del árbol

Las gimnospermas presentan un nexa evolutivo entre helechos (plantas sin semillas) y angiospermas (plantas con flores). Las coníferas incluyen a los árboles más antiguos y más grandes del mundo y constituyen el tipo de vegetación dominante en muchos ecosistemas (Brodrribb *et al.*, 2014). El clado contiene 6 familias con 68 géneros y muy pocas especies vivas (630 especies) en comparación con unas 250.000 especies conocidas de angiospermas (Wu and Chaw, 2014). Los pinos son árboles perennes que dominan ecosistemas con temperaturas que alcanzan desde -50°C en invierno en Noruega hasta por encima de 40°C en regiones subtropicales. Estos árboles son tolerantes a muchos factores de estrés debido a la estructura del xilema que es menos sensible a la cavitación. Pueden vivir en suelos pobres y hábitats problemáticos (Lev-yadun and Sederoff, 2000).

Pinus pinaster Aiton es una de las especies más comunes en la zona mediterránea occidental (Figura 1.6), mayoritariamente en la Península Ibérica y en España se extiende por una superficie de unos 2,4 millones de hectáreas (Alía *et al.*, 2009). Son árboles que alcanzan una altura alrededor de 20 metros, aunque existen excepciones que llegan incluso hasta 40 metros. Presentan una copa amplia que sólo ocupa aproximadamente el tercio superior del árbol. Tienen una raíz principal profunda y raíces secundarias muy desarrolladas que producen un enraizamiento muy potente. Crecen rápidos y tienen una longevidad entre 100 y 300 años. Entre las diferentes poblaciones se observa mucha variación en la distribución de biomasa dentro del árbol, resultado de un gran potencial de diferenciación adaptativa a las condiciones locales de la localización geográfica (Sanchez-Gomez *et al.*, 2010). Es una especie de gran importancia en su área de distribución para el medio ambiente y sus habitantes.

1.6.2 Valor ecológico y económico de coníferas

La madera producida por árboles forestales durante el crecimiento secundario representa un secuestro duradero del carbono fijado en la fotosíntesis. Junto con los océanos tienen un importante papel en el ciclo de carbono como mediadores del nivel atmosférico de CO₂ y además evitan la erosión del suelo en gran parte del planeta (Mackay *et al.*, 2012). El calentamiento global aumenta la frecuencia de incendios y plagas de insectos en los bosques boreales (Flannigan *et al.*, 2005; Kurz *et al.*, 2008; Shvidenko, 2011). Además, por la actividad humana grandes cantidades de nitrógeno se depositan como resultado del uso de carburantes fósiles, uso masivo de fertilizantes y altos niveles de contaminantes atmosféricas. En Europa se está observando la reducción de los bosques de coníferas y los árboles presentan síntomas como la pérdida de acículas, ramas colgantes y clorosis, probablemente debido al exceso de amonio (Galiano *et al.*, 2010). La pérdida en biomasa de los bosques conlleva una disminución en hábitats únicos, en agua limpia y también en recursos obtenidos por el turismo (Lucas *et al.*, 2013).

Aparte del valor ecológico, las coníferas como árboles leñosos tienen gran valor económico. A nivel mundial los tejidos leñosos de los árboles ganan en interés como fuente de bioenergía en la producción de etanol, sintegas, otros carburantes y productos químicos (Sannigrahi and Ragauskas, 2010). Además, las coníferas tienen un gran interés económico por su valor en madera y otros productos aprovechables. La madera de la especie *Pinus pinaster* en concreto es aplicada en la carpintería, construcción, tarima, tableros de fibra o partículas, etcétera. Es el árbol más usado para la madera en España con una producción anual de 3,7 millones de m³ (incluso corteza) que representó el 53,82 % de la producción total de cortas anuales de coníferas en el año 2012. La extracción de resina de esta especie está en auge con una producción anual que ha evolucionado de 1.821 toneladas en 2010 a 3.959 en 2011 hasta casi 7.000 toneladas en 2012 (Ministerio de agricultura, 2012).



Figura 1.6 Mapa de distribución de las poblaciones de *Pinus pinaster* Aiton. Las poblaciones de la especie, distribuidas en la región mediterránea, están marcadas con color oscuro sobre el mapa de la región.

1.6.3 Interés de la especie *Pinus pinaster* Aiton como modelo

Los pinos forman un modelo para la estructura y biosíntesis de madera en una conífera. El xilema en diferenciación es una fuente abundante de tejido metabólicamente activo, especializado en la formación de la pared celular secundaria y muerte celular programada (Lev-yadun and Sederoff, 2000). Las coníferas, al contrario de muchas otras plantas terrestres, prefieren al amonio como fuente de nitrógeno por encima del nitrato y su incorporación a través de las raíces es favorecida por la asociación con micorrizas (Kronzucker *et al.*, 1997). Como árboles perennes también son un modelo único para estudiar la asimilación de amonio ya que, en muchas especies, el N asimilado del suelo no

es usado inmediatamente pero es almacenado en la corteza para ser movilizado durante la siguiente estación de crecimiento para el nuevo desarrollo de órganos (Coleman *et al.*, 1994).

La secuenciación de especies gimnospermas es tedioso debido a su enorme genoma con un tamaño medio de unos 25 Gbp (Morse *et al.*, 2009), pero especialmente por la acumulación de secuencias repetitivas probablemente debido a retrotransposones (Kamm *et al.*, 1996; Islam-faridi *et al.*, 2009) y gran cantidad de pseudogenes (Mackay *et al.*, 2012). La especie *Pinus pinaster* es el modelo más avanzado en Europa en genómica de coníferas cuyo transcriptoma tiene un total de 210.513 unigenes secuenciados que han proporcionado un catálogo de 26.020 transcritos únicos con ortólogos en las bases de datos y 17.815 son secuencias cDNA de longitud completa (Canales *et al.*, 2014). Su genoma consiste en 12 pares de cromosomas grandes y metacéntricos y como todos los pinos son diploides. Típicamente las secuencias de DNA de los genes ortólogos son 60 a 80 % conservadas con angiospermas y alrededor del 10 % de los genes EST de pino no tienen ortólogos en angiospermas.

Las coníferas constituyen una fuente renovable de recursos que enriquecen la biosfera y la población humana (Lev-yadun and Sederoff, 2000). Como especie de gran importancia en la región mediterránea, un aumento en la productividad de *P. pinaster* impactaría significativamente en la generación de fuentes renovables bioenergéticas y productos de gran valor económico. Para alcanzar este objetivo es necesario entender el proceso y la regulación del crecimiento del árbol. Es una especie con una alta tasa de diferenciación genética, de polimorfismos y heterozigocidad en comparación con otras coníferas (Salvador *et al.*, 2000) lo que también ha contribuido a su interés como especie modelo y permite la identificación de genotipos de mayor interés.

2. ANTECEDENTES Y OBJETIVOS

2.1 Antecedentes

Como vehículo para el transporte y reserva de nitrógeno, la asparragina se relaciona con tejidos sumideros donde la principal vía de movilización del nitrógeno contenido en el grupo amido del aminoácido implica a la enzima asparraginasa (Lea *et al.* 2007). Aunque conocemos, en líneas generales, la implicación de la enzima en la movilización de nitrógeno en plantas, desconocemos los mecanismos moleculares que coordinan y controlan espacial y temporalmente su actividad. La mayoría de los estudios se han realizado en plántulas herbáceas anuales, mientras que en plantas leñosas perennes son muy limitados. Los árboles son distintos de las plantas herbáceas en algunos aspectos particulares de su fisiología y desarrollo, que resultan de gran importancia para la producción de biomasa (Cantón *et al.*, 2005). Durante el desarrollo y crecimiento del tronco, afrontan un consumo masivo de carbono y nitrógeno para proveer la síntesis de celulosa y lignina, proceso que debe estar perfectamente controlado para evitar descompensaciones metabólicas en el árbol. Dado que la disponibilidad de nitrógeno puede ser un factor determinante de la producción y calidad de la madera de árboles (Novaes *et al.*, 2009) consideramos que la actividad del gen de la asparraginasa en su función de movilización del elemento podría ser de gran relevancia para la producción de biomasa. Mediante experimentos de hibridación *in situ*, se comprobó que la expresión de una asparraginasa de pino está confinada a las células de la región del cambium por lo que podría jugar un papel importante en el desarrollo vascular (Cañas *et al.*, 2007). Por otro lado, la caracterización bioquímica de asparraginasas de especies angiospermas expresadas en células de *E. coli*, mostraron la capacidad de autoprocesamiento espontáneo *in vitro* (Bruneau *et al.*, 2006). Sin embargo, se comprobó que la asparraginasa de pino recombinante no presentaba esta capacidad de autoprocesamiento

espontáneo (Cañas *et al.*, 2007), por lo que se desconocía si en coníferas el precursor requiere de algún factor para poder ser procesado a la forma activa.

2.2 Objetivos

Considerando los antecedentes expuestos se plantearon como objetivos para este trabajo:

- Identificar las proteínas con actividad asparraginasa en pino marítimo (*Pinus pinaster*) con la ayuda del transcriptoma disponible, y realizar una caracterización básica que nos permita deducir su función en plántulas.
- Caracterización bioquímica más detallada de las que tengan una mayor relevancia en la movilización de N.
- Análisis transcripcional de los correspondientes genes.

3 MATERIAL Y MÉTODOS

3.1 Material biológico

Las cepas de *E. coli* DH5alfa y BL21-AI (Invitrogen Corporation, Carlsbad, CA, EE.UU.) se emplearon para la propagación del plásmido y la expresión de la proteína recombinante, respectivamente, en un sistema procariota. La cepa C58C1 de *Agrobacterium tumefaciens* se empleó para infiltrar hojas de plantas de *Nicotiana benthamiana* e inducir la expresión de la proteína recombinante *in planta*.

Las plantas de *N. benthamiana* se obtuvieron mediante siembra de semillas en sustrato para horticultura (Pindstrup professional, Terreau) humedecido con agua destilada. Las semillas de *P. pinaster*, procedentes de Sierra Bermeja (Estepona, España), se embebieron durante 24 horas en agua destilada con aireación y seguidamente se sembraron en vermiculita. Todas las plantas se germinaron a 22°C en luz y 18°C en oscuridad en fotoperiodo de 16 h luz/8 h oscuridad, regando regularmente con agua destilada.

Las plántulas de *Pinus pinaster* se muestrearon eliminando los cotiledones primero y luego separando hipocótilos y radículas. Como las plantas de esta especie no están sincronizadas en la germinación, la clasificación del estadio de desarrollo se basó en la longitud de los cotiledones, la aparición de acículas y el aspecto del tallo que pasa de color verde a marrón, coincidiendo con el aumento de lignificación del mismo (ver tabla 3.1 y figura 3.1). Las muestras de radícula recién germinada se recolectaron en el estadio E0 y se separaron las radículas que tenían una longitud de 1 cm. Para el análisis de distintos segmentos de las radículas, las muestras de este órgano se obtuvieron cuando las plantas llegaron a un estadio E4, dividiéndose en fracciones de 1 cm, enumeradas desde el ápice (R1) hasta 3 o 4 partes según la longitud que tenía la radícula. El muestreo de plántulas jóvenes se hizo siempre sobre un conjunto de plántulas crecidas en las mismas condiciones.

Tabla 3.1 Clasificación de los estadios de desarrollo de plántulas de pino.

Estadio de desarrollo	Descripción morfológica del criterio de clasificación de las plántulas	Promedio de días tras imbibición
E0	La radícula es el único órgano que emerge de la cáscara	7
E1	Los cotiledones, todavía envueltos por la cubierta seminal, tienen una longitud de 0,5 cm	10
E2	Los cotiledones miden 1,5 cm y mantienen contacto con el tegumento de la semilla	14
E3	Los cotiledones miden 2,5 cm y la cubierta seminal empieza a desprenderse	20
E4	Los cotiledones miden 3 cm y la cubierta seminal se ha desprendido	25
E5	Aparecen las primeras acículas y se observa lignificación gradual del hipocótilo	30
E6	El hipocótilo está lignificado en su totalidad	40

El análisis de la expresión a lo largo del tallo se realizó en muestras de 3 plantas individuales de *P. pinaster* procedentes de la Sierra de Gredos (España) con una edad de 2 años. Se separó el tallo de la raíz y se eliminaron todas las acículas, dividiéndose posteriormente cada tallo en tres partes según su morfología: la parte superior del tallo de nuevo crecimiento y de color verde, que denominamos “brote”, la parte intermedia con un grado de lignificación intermedio “tallo superior” y la parte inferior con un mayor grado de lignificación y ausencia de acículas “tallo inferior”.

Las 11 poblaciones de pino marítimo sobre las cuáles se comparó la región promotora del gen *PpASPG1* se distribuyen entre Francia, España y Marruecos y figuran en el mapa de la figura 3.2. Las plantas tenían una edad aproximada de 2 años y se recolectaron las acículas para la extracción de DNA.

Las muestras de tejidos de árboles de *P. pinaster* adultos se obtuvieron a partir de una población natural situada en la Sierra Bermeja, Málaga, tal y como se describió previamente (Pacheco-Villalobos, 2008; Díaz-Moreno, 2010).

Todos los tejidos se congelaron inmediatamente después de la recolección en nitrógeno líquido y se guardaron a -80°C hasta su uso.

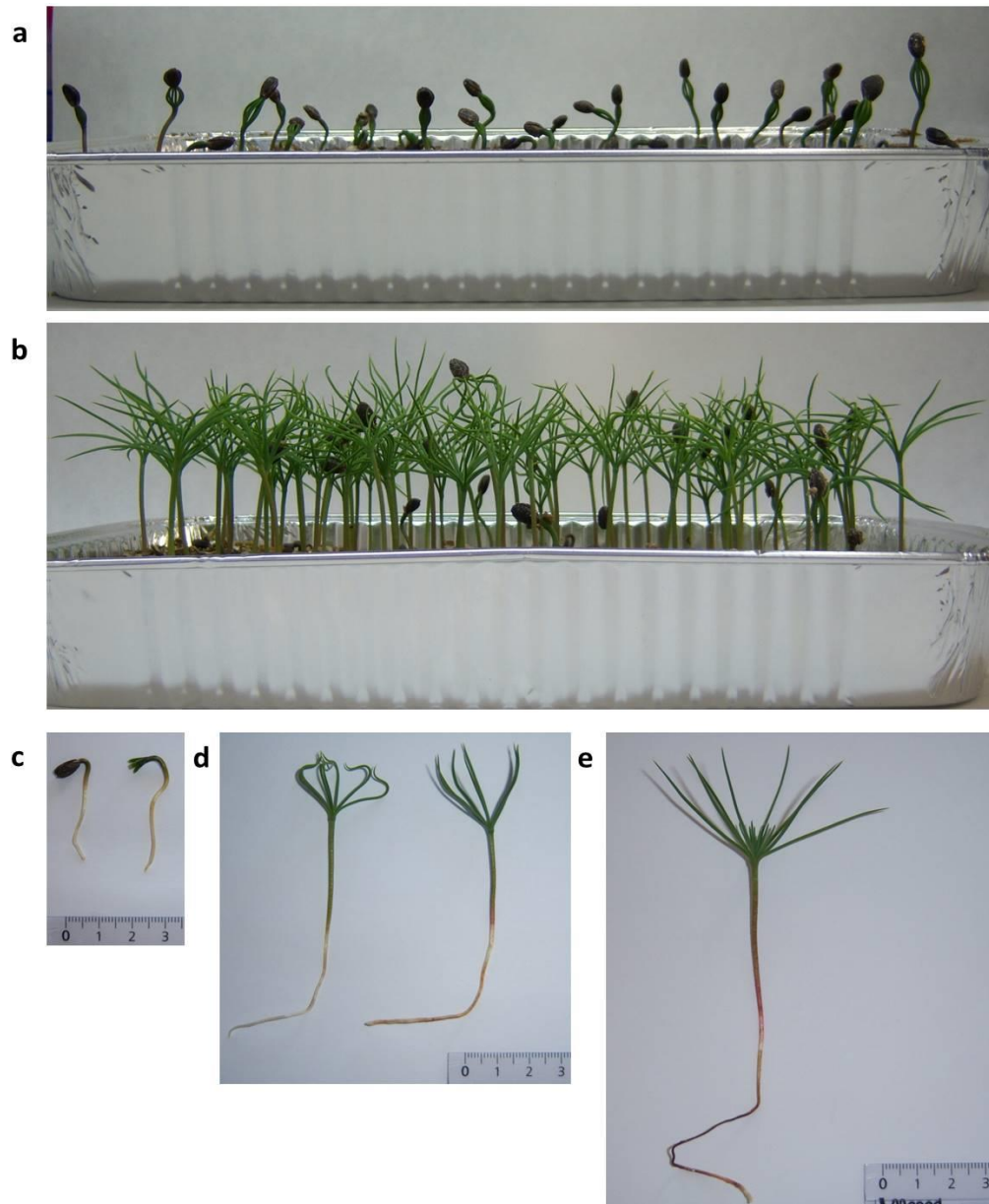


Figura 3.1 Crecimiento de plántulas de *P. pinaster*. Plántulas germinadas 2 semanas después de la imbibición de las semillas (a) y 3 semanas (b). Como la germinación de las semillas de esta especie no está sincronizada, se determinan los estadios de desarrollo según la morfología. Por ejemplo estadio E1 (c), E3 (d) o E5 (e).

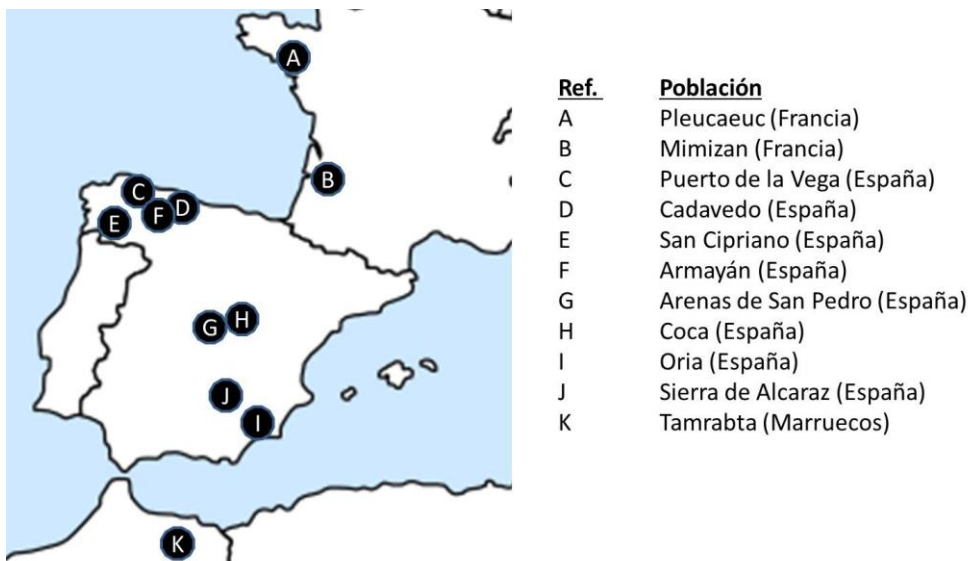


Figura 3.2 Mapa con la distribución de 11 poblaciones de *Pinus pinaster* incluidas en el análisis de variabilidad de secuencia en la región promotora.

3.2 Condiciones experimentales de crecimiento de las plantas

Durante las primeras semanas del crecimiento de plántulas de pino el megagametofito aporta nutrientes al embrión. Para evaluar el efecto de una fuente de amonio externa sobre la expresión del gen ASPG usamos plántulas en estadio E4 como se ha descrito en la sección 3.1, regadas solamente con agua destilada y de las cuales se había desprendido los restos de la semilla, para asegurarnos que no quedaban reservas del megagametofito. Luego se regaron con una solución nutritiva con la siguiente composición: Na_2HPO_4 20 mg/L, KCl 40 mg/L, CaCl_2 40 mg/L, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 20 mg/L, Fe quelado 5,6 mg/L, $\text{NaMoO}_4 \cdot 7\text{H}_2\text{O}$ 0,001 mg/L, $\text{CuSO}_4 \cdot 7\text{H}_2\text{O}$ 0,006 mg/L, H_3BO_3 0,009 mg/L, $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 0,1 mg/L y $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ 0,7 mg/L. Las plántulas se dividieron en tres conjuntos y se regaron 1 vez por semana con la solución nutritiva suplementada con sulfato de amonio como única fuente de nitrógeno a una concentración final de 3 mM, 10 mM o su volumen en agua estéril para las plántulas que no disponían de ninguna fuente de nitrógeno. Se recolectaron

muestras de cotiledones, hipocótilos y radículas del 50% de las plántulas a 24 horas después del primer riego y las plántulas restantes se siguieron regando con la solución un total de 5 veces para recolectar las muestras a 24 horas después del último riego (29 días después del primer riego).

Para evaluar la expresión del gen en condiciones de estrés se usaron grupos de plántulas jóvenes de *P. pinaster* y se les aplicaron tratamientos simulando diferentes factores de estrés como describieron Pascual *et al.* (2015). El megagametofito fue cuidadosamente eliminado y se seleccionaron las plántulas en estadio E1, las cuales se pasaron a macetas con vermiculita a una densidad de 12 individuos/maceta y se regaron con agua destilada. Cinco días después se aplicaron los tratamientos mientras que un grupo se siguió regando con agua destilada sin ningún tratamiento como grupo control. Se regó con NaCl 250 mM disuelto en el agua destilada para la condición de estrés salino. Las fitohormonas se diluyeron en agua bidestilada estéril a 50 μ M (ácido abscísico) y 100 μ M (ácido jasmónico) y la solución se pulverizó sobre los cotiledones. Para el tratamiento "frio" se pasaron las macetas a 4°C y las heridas se provocaron quitando un tercio de los cotiledones y pinchando los hipocótilos.

3.3 Extracción de RNA y síntesis de cDNA

Las muestras de tejidos se pulverizaron con la ayuda de nitrógeno líquido en un mortero hasta obtener un polvo fino y se extrajo RNA total a partir de 100 mg mediante un protocolo basado en el descrito por Liao *et al.* (2004) con algunas modificaciones (Canales *et al.*, 2012). Se mezcló el polvo con 650 μ L tampón de extracción (PVP40 2 % (p/v), CTAB 3 % (p/v), Tris-HCl 100 mM pH 8, EDTA 30 mM pH8, NaCl 2 M, 2-mercaptoetanol al 4 % (v/v)). A continuación se realizaron dos extracciones con 650 μ L de una mezcla cloroformo:alcohol isoamílico (24:1) y se separaron las fases por centrifugación a 16.090 x g durante 5 min a 4°C. Enseguida se precipitó el RNA por adición de LiCl a una concentración final de 3 M y se incubó durante una hora en hielo. El RNA se

recuperó mediante centrifugación a 18.660 x g durante 25 min a 4°C. Para eliminar los restos de sales que co-precipitaron con el RNA se añadió al pellet 500 µL etanol 70 % (v/v) frío y se centrifugó otros 5 min a 18.660 x g para luego resuspender el pellet en 40 µL de agua bidestilada estéril. Para eliminar DNA contaminante se incubó durante 20 min en un baño a 37°C con 5 U de DNasa RQ1 libre de RNasa (Promega Corp, Madison, WI, EE.UU.). Tras lo cual se purificó el RNA empleando el sistema comercial "Nucleospin PCR Clean up kit" (Macherey-Nagel GmbH, Düren, Alemania) para eliminar la DNasa y otros posibles inhibidores de la síntesis de cDNA, siguiendo las instrucciones del manual con las siguientes modificaciones: se añadió 300 µL de etanol 100% a 300 µL de "binding buffer" antes de cargar la columna y la elución del RNA puro se realizó con 20 µL de agua bidestilada estéril. La cantidad de RNA fue determinado en un NanoDrop® NDcastro 1000 spectrophotometer (NanoDrop Technologies, Inc. Wilmington, EE.UU.) y la integridad y pureza del RNA se evaluó mediante espectrofotometría y electroforesis en gel desnaturizante de agarosa 1 % (p/v) con formaldehído 6,7 % (p/v). Como tampón de electroforesis se utilizó MOPS 20 mM pH 7, Na₂EDTA 1 mM y NaAc 5 mM.

El cDNA se sintetizó usando los reactivos del sistema comercial "iScript cDNA Synthesis kit" (Bio-Rad, Hercules, CA, EE.UU.) que incluye una transcriptasa inversa con actividad RNasa H⁺ asociada. Se mezcló 1 µg del RNA con 1 µL de iScript reversotranscriptasa, 4 µL de la mezcla de reacción iScript 5X que incluye los cebadores oligo(dT) y hexaméricos de secuencia aleatoria, y agua estéril libre de RNasa hasta un volumen final de 20 µL. La mezcla se incubó 5 min a 25°C, 30 min a 42°C y 5 min a 85°C. Posteriormente el cDNA sintetizado se diluyó a una concentración equivalente a 2,5 ng/µL, 5 ng/µL o 12,5 ng/µL de RNA según conviniera para su uso y se almacenó a -20°C. Este método de síntesis de cDNA se aplicó para el aislamiento de las secuencias de longitud completa de cDNA de los genes *PpASPG2* y *PpASPG3* y para los análisis de niveles de mRNA mediante qRT-PCR. La región codificante de *PpASPG1* se aisló sobre cDNA sintetizado mediante el método RLM RACE como se explica en el párrafo siguiente.

3.4 Clonación de las secuencias completas de cDNA de los genes de asparraginasas en pino

Se aplicó una estrategia de amplificación del extremo 5' de cDNA mediada por RNA ligasa usando el kit "FirstChoice RLM RACE" (Applied Biosystems/Ambion, Austin, TX) para obtener la secuencia de cDNA completa del gen *PpASPG1* de *P. pinaster*, siguiendo las instrucciones del manual del kit. Se partió de 10 µg de RNA total extraído de xilema de árboles adultos el cual fue tratado con una fosfatasa alcalina bovina (CIP) para eliminar los grupos PO₄ en los extremos 5' de las moléculas de RNA ribosomal, tRNA, restos de DNA genómico y de fragmentos incompletos de mRNA. Las moléculas de mRNA completas no se ven afectadas por el tratamiento debido a la presencia de una 7-metilguanosina (CAP) en su extremo 5'. Seguidamente la muestra fue tratada con una pirofosfatasa ácida de tabaco (TAP) para eliminar el CAP. De tal manera los mRNA completos resultan las únicas moléculas en la muestra que exponen un grupo fosfato en el extremo 5' y al cual, en el siguiente paso, se unen unos adaptadores de secuencia conocida en una reacción catalizada por la T4 RNA ligasa. A continuación se realizó una reversotranscripción del RNA empleando cebadores decaméricos de secuencia aleatoria para obtener cDNA que contiene el inicio de la transcripción.

Se diseñaron cebadores A1-1 y A1-2 anidados en sentido hacia 5' (las secuencias de los cebadores figuran en el suplemento 2) sobre una secuencia EST previamente aislada (Villalobos *et al.*, 2012) para amplificar la región 5'UTR sobre el molde 5'RACE cDNA. El uso de estos oligonucleótidos específicos en combinación con los cebadores del kit que se unen a la secuencia del adaptador en dos reacciones de PCR consecutivos asegura la especificidad del fragmento amplificado. La primera reacción de PCR con los cebadores externos se hizo en un volumen total de 50 uL con 1 uL del cDNA, dNTPs a 0,2 mM, los cebadores a 0,4 µM y 2,5 U de "Pfu Turbo DNA polimerasa" (Stratagene, Agilent technologies, CA) en el tampón proporcionado por la casa comercial diluida a concentración 1x. Las

condiciones de la PCR fueran las siguientes: 2 min a 95°C de desnaturalización inicial seguida por 30 ciclos de 30 s a 95°C, 30s a 60°C y 1 min 15s a 72°C, con un paso final de 10 min a 72°C. Luego se utilizaron 2 µL del producto como molde en la segunda PCR que se desarrolló en las mismas condiciones con los cebadores internos.

El producto final del 5'RACE se analizó mediante electroforesis en un gel de agarosa al 0,8 % (p/v) y se purificó mediante el sistema comercial "Nucleospin PCR Clean up kit" (Macherey-Nagel GmbH, Düren, Alemania). El fragmento purificado se clonó en el vector pBluescript SK+ (Agilent Technologies, Santa Clara, CA, EE.UU.) previamente digerido con la enzima de restricción *EcoRV* (Roche applied science, Mannheim, Alemania) poniendo una relación inserto:vector 5:1. Para ello se añadió 1 µL (2U) de la enzima T4 DNA ligasa (Promega Corp, Madison, WI, EE.UU.), 1 µL de su tampón a 10x y agua estéril hasta un volumen final de 10 µL. La mezcla se incubó a temperatura ambiente durante 16 horas y después se transformaron células competentes de la cepa DH5α de *E. coli*. Se mezcló el volumen total de la ligación con 100 µL de células bacterianas y se incubaron 10 min en hielo. Enseguida se pasó el vial a un baño a temperatura de 42°C exactamente 1 min para posteriormente devolverlo al hielo. 2 min después se la añadió 950 µL de medio Luria-Bertani (LB) y se incubó 1 hora a 37°C con agitación a 200 rpm para después distribuir en placas las células transformadas en un medio selectivo de LB y agar al 1,5% (p/v) con ampicilina 100 µg/mL, IPTG 0,2 mM y X-gal a 0,004 % (p/v). Después de una noche a 37°C se seleccionaron algunas colonias positivas de color blanco las cuales se cultivaron en 4 mL de LB líquido con el antibiótico durante otra noche a 37°C con agitación. El día siguiente se comprobó la presencia y el tamaño del inserto mediante PCR con los cebadores T3 y T7 sobre 2,5 µL del cultivo bacteriano y se almacenaron alícuotas de las colonias positivas en -80°C en presencia de glicerol 15% (v/v) como agente crioprotector.

El plásmido se aisló mediante minipreparación de DNA plásmidico por lisis alcalina y el inserto se secuenció en un secuenciador Beckman Coulter

CEQ8000. Sobre la secuencia obtenida se diseñó el cebador A1-3 al 5'UTR para poder aislar la secuencia completa del cDNA de *PpASPG1* mediante la técnica de 3'RACE. Para ello se sintetizó cDNA 3'RACE a partir de 1 µg de RNA por retrotranscripción usando el cebador oligo-dT que tiene unido un adaptador con secuencia conocida, proporcionado por el sistema comercial. La secuencia completa *PpASPG1* se amplificó por combinación del cebador A1-3 específico con el 3'RACE externo del kit en una PCR con las mismas condiciones como la amplificación 5'RACE con la excepción del paso de hibridación que se hizo a una temperatura de 59 °C. El producto amplificado se clonó en *E. coli* para su propagación y posterior secuenciación de igual manera como se ha descrito para 5'RACE. Se hicieron un total de tres reacciones de PCR independientes y las secuencias obtenidas se compararon para deducir la secuencia del cDNA de longitud completa libre de errores.

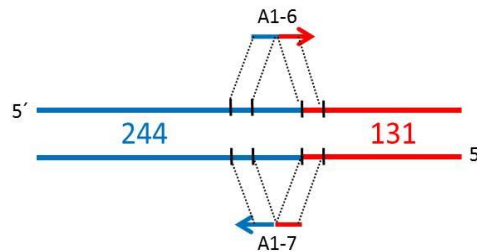
Las secuencias codificantes de los genes *PpASPG2* y *PpASPG3* se obtuvieron con los cebadores A2-1 con A2-2 para *PpASPG2* y A3-1 con A3-2 para *PpASPG3*, los cuales se habían diseñado sobre las regiones UTR de los unigenes 15077 y 36994 respectivamente que están disponibles en la base de datos Sustainpine V.3.0. (http://www.scbi.uma.es/sustainpinedb/home_page). Para ello se usó como molde 50 ng de cDNA de plántulas de pino jóvenes sintetizado como se ha descrito en la sección 3.3. La PCR se desarrolló con el siguiente programa de temperaturas: 2 min a 95°C de desnaturalización inicial seguida por 30 ciclos de 30 s a 98°C, 30s a 60°C para *PpASPG2* / 57°C para *PpASPG3* y 1 min a 72°C con un paso final de 10 min a 72°C. Las reacciones de PCR se hicieron con 1 U de una polimerasa con una baja tasa de error "Iproof High-Fidelity DNA polymerase" (Bio-Rad, Hercules, CA, EE.UU.) en un volumen de 50 uL con dNTPs a 0,2 mM, los cebadores a 0,5 µM y el tampón proporcionado por la casa comercial diluida a concentración 1x. La purificación, clonación y secuenciación de los fragmentos de PCR se realizaron tal y como se ha descrito para el gen *PpASPG1*.

3.5 Construcción de vectores para la expresión de la proteína recombinante en hojas de *N. benthamiana*

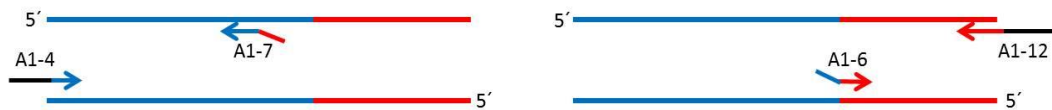
Los cebadores A1-4 y A1-5 se usaron para amplificar el ORF del gen *PpASPG1* usando como molde el plásmido conteniendo el fragmento del transcrito completo previamente aislado, incorporando un único sitio de restricción para *PacI* en el extremo 5' y un sitio *NotI* en 3'. Después de la digestión con las respectivas enzimas de restricción (Fast Digest, Thermo Fisher Scientific Inc, Waltham, MA, EE.UU.), se ligó en el vector pJL-TRBO (Lindbo, 2007), previamente linearizado por digestión con las mismas enzimas de restricción, para la expresión *in planta*. La construcción se introdujo en la cepa DH5α *E. coli* para propagación y verificación de la secuencia antes de su subclonación en *A. tumefaciens* C58C1 mediante transformación por electroporación.

Los ORF de los genes *PpASPG2* y *PpASPG3* se amplificaron sobre las minipreparaciones de los correspondientes plásmidos. Para lo cual se emplearon los cebadores A2-3 con A2-4 y A3-3 con A3-4, respectivamente, que incorporan los sitios de restricción *PacI* en el extremo 5' y *NotI* además de una etiqueta "tag" HA y poli-His en el extremo 3' de la secuencia. Para el gen *PpASPG1* se hizo una construcción incorporando el mismo tag mediante cebadores A1-4 en combinación con A1-12.

Diseño de 2 cebadores complementarios que flanquean la región a eliminar



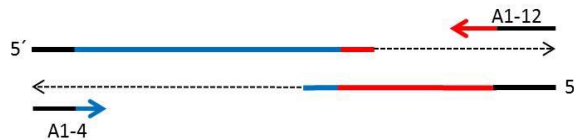
Amplificación de las subunidades por separado en dos reacciones de PCR, en combinación con cebadores externos que incorporan sitios de restricción



Productos de las dos reacciones de PCR:



Combinación de los productos como molde en una tercera PCR, con sólo los cebadores externos



Producto final:

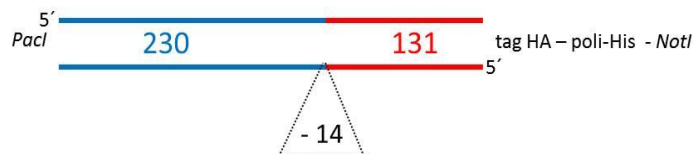


Figura 3.3 Procedimiento para la construcción de variantes de la proteína PpASPG1 que presentan eliminaciones graduales en la RV. Se representan los cebadores usados para la amplificación de PpASPG1-D1 como ejemplo. Las líneas azules representan la doble hebra de DNA que codifica para residuos de la subunidad α , las rojas de la subunidad β y el número de aminoácidos figura entre las líneas. El total de 3 reacciones de PCR produce un fragmento que codifica para el precursor con la eliminación de 14 aminoácidos de la RV. Se ha usado el mismo procedimiento para las otras variantes con diferentes parejas de cebadores complementarios para eliminar regiones de mayor número de residuos.

El procedimiento seguido para producir la proteína PpASPG1 con distintas eliminaciones en el extremo carboxilo de la subunidad α se muestra en la figura 3.3. Se diseñaron juegos de 2 oligonucleótidos complementarios, la secuencia de estos oligos incluía 15 pb de cada uno de los extremos que flanquea la secuencia a eliminar y que en la construcción final quedarían fusionados (oligos A1-6 y A1-7 para PpASPG1-D1), los cuales se combinaron en reacciones de PCR con los oligonucleótidos A1-12 y A1-4 respectivamente para amplificar las regiones codificantes de ambas subunidades por separado. Luego los productos de las primeras reacciones de PCR se combinaron como molde en una tercera reacción de PCR con los cebadores externos A1-12 y A1-4, juntando los fragmentos en una secuencia mutada con la eliminación y presentando un tag HA y poli-His a su extremo 3'. Las formas PpASPG1-D2 y PpASPG1-D3 se generaron siguiendo el mismo procedimiento, empleando los oligonucleótidos A1-8 y A1-9 para PpASPG1-D2 y A1-10 y A1-11 para PpASPG1-D3, en combinación con A1-12 y A1-4.

3.6 Expresión *in planta* y extracción de proteínas recombinantes

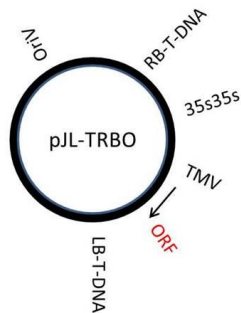
El proceso completo usado para la expresión de las proteínas recombinantes *in planta* se resume en el lado izquierdo de la figura 3.4. Se transformaron células de *A. tumefaciens* con los vectores de expresión pJL-TRBO mediante electroporación y se sembraron en un medio selectivo LB con adición de los antibióticos kanamicina 50 ng/ μ L, rifampicina 50 ng/ μ L y tetraciclina 5 ng/ μ L. Las colonias positivas se picaron y cultivaron durante 24 horas a 28°C con agitación a 200 rpm. El día siguiente las células se recolectaron por centrifugación durante 10 min a 2.250 x g y se lavaron con una solución de MES 52 mM pH 5,6, glucosa 0,5 % (p/v), NaH₂PO₄ 2,1 mM, NH₄Cl 18,5 mM, MgSO₄·7H₂O 1,2 mM, KCl 2 mM, CaCl₂ 0,09 mM, FeSO₄·7H₂O 0,09 mM y acetosiringona 0,2 mM, tras lo cual se resuspendieron otra vez en el mismo medio para dejarlas incubar durante 4 horas. Después se recolectaron las células de nuevo por centrifugación para resuspenderlas en una solución de MES 10 mM pH 5,6 con acetosiringona 0,2 mM hasta una densidad óptica a 600 nm de 0,5. La inoculación de hojas de plantas de *N. benthamiana* de 6 semanas de edad se realizó tocando el envés con una aguja para generar una herida pequeña y después infiltrando el cultivo de *A. tumefaciens* en el lugar de la herida con la ayuda de una jeringa. Las zonas de las hojas infiltradas se marcaron con rotulador para después de 5 días poder recolectar el tejido infiltrado que expresa las proteínas recombinantes. Las muestras se homogenizaron en un mortero con nitrógeno líquido y al polvo macerado se añadió el tampón de extracción (HEPES 50 mM pH 7.8, KCl 50 mM, EDTA 5 mM, glicerol 10% (v/v), 2-mercaptoethanol 10 mM, y PMSF 1 mM) en una relación tampón:materia vegetal 2:1 (v:p). Las proteínas solubles se separaron de la fracción insoluble por centrifugación a 18.660 x g a 4°C durante 20 min y el sobrenadante se guardó a -80°C hasta su uso.

Expresión de proteínas recombinantes

- Aislamiento de región codificante por PCR

- Ligación en pBSK
- Transformación de *E. coli*
 - Propagación de pBSK
 - Secuenciación del inserto

- Subclonación en pJL-TRBO



- Transformación de *E. coli*
 - Propagación del plásmido
 - Secuenciación de la construcción
- Transformación de *A. tumefaciens*
- Agroinfiltración de *N. benthamiana*



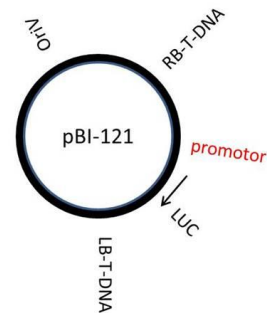
- Extracción de proteínas totales

- Purificación de la proteína recombinante
 - caracterización

Análisis funcional de la región promotora

- Aislamiento de región promotora por PCR

- Subclonación en pBi121



- Cuantificación de actividad luciferasa

Figura 3.4 Flujo de trabajo para la expresión de proteínas recombinantes *in planta* y el análisis funcional de la región promotora del gen *PpASPG1*. Se usó el método de agroinfiltración de hojas de *N. benthamiana* tanto para producir las proteínas recombinantes de ASPG de pino *in planta* (texto a la izquierda) como para comprobar la funcionalidad de la región promotora (texto a la derecha). Los pasos en común entre ambos procesos están escritos en color azul en medio. Esquemas simplificados de los vectores usados marcan en color rojo la posición y función del inserto. OriV: origen de replicación vegetativo, ORF: secuencia codificante para la proteína, TMV: secuencia del Virus del Mosaico del Tabaco, 35s35s: promotor 35s de CaMV por duplicado, LUC: secuencia codificante para luciferasa. El T-DNA engloba el fragmento entre los bordes RB-T-DNA y LB-T-DNA.

3.7 Purificación de proteínas recombinantes de pino, expresadas en *N. benthamiana*

Para la purificación de las proteínas recombinantes PpASPG1 y PpASPG1-D3 se aplicó un protocolo convencional como descrito por Sodek y Lea (1993) con algunas modificaciones. Se precipitaron las proteínas por la adición gradual de $(\text{NH}_4)_2\text{SO}_4$ hasta un 80 % (p/v) de saturación seguido por una centrifugación durante 10 min a $18.660 \times g$ a 4°C . El precipitado se resuspendió en medio de extracción y se le añadió gradualmente 1,5 volúmenes de EtOH hasta obtener una concentración final de 60% (v/v). La mezcla se incubó durante 1 hora en hielo y luego se separaron las proteínas solubles incluida la asparraginas, de las insolubles por centrifugación durante 20 min a $2.000 \times g$ y 4°C . El sobrenadante fue dializado extensivamente en el tampón de extracción para eliminar el EtOH. Antes de pasar a la cromatografía de intercambio iónico se incluyó un paso de concentración por diálisis inversa en sacarosa y la proteína concentrada fue cargado sobre una columna de intercambio iónico DEAE-Sephacel (GE Healthcare Bio-Sciences AB, Uppsala, Suecia) con un volumen de 50 mL previamente equilibrada con un tampón HEPES 50 mM pH 7,8, KCl 50 mM, EDTA 5 mM, glicerol 10% (v/v) y 2-mercaptoethanol 10 mM. Se lavó con un volumen del mismo tampón equivalente a dos volúmenes de la columna y la enzima se eluyó por la aplicación de un gradiente lineal de KCl (50-500 mM), todo a un flujo de 42 mL/h y se recolectaron fracciones de 3 mL. Las fracciones con actividad ASPG se juntaron, se dializaron extensivamente con el tampón de extracción y se concentraron por diálisis inversa en sacarosa. Después se cargó la muestra en una columna de cromatografía de filtración en gel Sephacryl S-300 (Sigma-Aldrich Co. LLC. Saint Louis, MO, EE.UU.) con un tamaño de 93×2 cm previamente equilibrada con un tampón HEPES 50 mM pH 7,8, KCl 100 mM, EDTA 5 mM. Se recolectaron fracciones de 1 mL a una velocidad de 16 mL/h y las fracciones que presentaban actividad asparraginas se juntaron y concentraron. Durante todo el proceso de purificación las muestras se mantuvieron a una temperatura de 4°C . La cuantificación de las

proteínas se llevó a cabo por el método descrito por Bradford (1976) usando albumina bovina comercial como estándar.

3.8 Construcción del vector para la expresión de la proteína recombinante en *E. coli*

Para expresar la proteína recombinante en *E. coli* se aplicó el método "Gateway technology" (Invitrogen Corporation, Carlsbad, CA, EE.UU.) partiendo de 100 ng del plásmido con la secuencia completa del gen *PpASPG1* como molde. En una primera PCR con los oligonucleótidos A1-15 y A1-16 se amplificó la región codificante para la proteína en un volumen de 25 μ L con los cebadores a 0,4 μ M, dNTPs a 0,2 mM cada uno, 1 U de "Iproof High-Fidelity DNA polymerase" en el tampón proporcionado, diluido a concentración 1x (Bio-Rad, Hercules, CA, EE.UU.) y empleando el siguiente programa de temperaturas: 1 min a 98°C de desnaturalización inicial seguido por 20 ciclos de 10 s a 98°C, 30s a 62°C y 1 min a 72°C y un paso final de 10 min a 72°C. El volumen total fue 25 μ L. El producto de la primera PCR se cargó en un gel de agarosa y se purificó después de la electroforesis con el sistema comercial "Nucleospin PCR Clean up kit" (Macherey-Nagel GmbH, Düren, Alemania). Luego se usaron 2 ng como molde en una segunda reacción de PCR con cebadores attB1 y attB2 que completan la secuencia del sitio de recombinación en los extremos del fragmento. El programa de PCR contenía los siguientes pasos: 1 min a 98°C, 5 ciclos de 10s a 98°C, 20s a 45°C y 1 min a 72°C, 20 ciclos de 10 s a 98°C, 30s a 55°C y 1 min a 72°C y un paso final de 10 min a 72°C.

El fragmento final de la segunda PCR purificado se insertó en el vector de entrada pDONR207 mediante recombinación, siguiendo las instrucciones del manual suministrado con el sistema comercial. Una cantidad de 200 ng del vector se mezclaron con 200 ng del inserto *PpASPG1* y 1 μ L de una mezcla de reacción que incluye la CLonasa BP, en un volumen total de 5 μ L. La mezcla se incubó durante 18 horas a 25 °C. Luego se transformaron células de *E. coli*

DH5α con la mezcla de recombinación para la propagación del vector de entrada y comprobación de la secuencia del inserto. Finalmente, se seleccionaron las colonias positivas que crecieron en un medio LB selectivo con gentamicina a 12,5 ng/μL.

Para la generación del vector de expresión, que incluyera la secuencia codificante para la PpASPG1 y un tag poli-His al extremo amino de la proteína recombinante, se desarrolló una reacción de recombinación con 190 ng del vector pDEST17, 380 ng del vector de entrada con el inserto y 1 μL de una mezcla de reacción con Clonasa LR en un volumen de 5 μL. La mezcla se incubó 18 horas a 25 °C para después transformar células de la cepa DH5α, las cuales se crecieron en LB con ampicilina a 100 μg/mL. A partir de las colonias obtenidas se realizaron cultivos líquidos para el aislamiento del plásmido, el cual se verificó mediante secuenciación. Se construyó también un vector para la expresión de la variante PpASPG1-D3, que carece de RV. Para ello se aplicó la misma estrategia de “Gateway Technology” y los mismos cebadores, usando como molde el plásmido con la secuencia PpASPG1-D3 que se había usado para su expresión *in planta*.

Después de haber comprobado por secuenciación que las construcciones en pDEST17 eran correctas, se transformaron células competentes de la cepa BL21-AI por choque térmico y las células transformadas se seleccionaron en medio LB suplementado con ampicilina a 100 μg/mL.

3.9 Inducción de la expresión y purificación de la proteína recombinante de *E. coli*

Siguiendo las instrucciones del manual de Invitrogen, se cultivaron las bacterias BL-21-AI transformadas con el vector de expresión en un medio de LB con ampicilina 100 μg/mL a 37°C con agitación hasta llegar a una densidad óptica a 600 nm de 0,5. Entonces, se añadió arabinosa al 0,2 % (p/v) y el

cultivo se incubó otras 6 horas, bajando la temperatura a 28 grados. Las células bacterianas crecidas se recolectaron por centrifugación durante 10 min a $2.750 \times g$ y 4°C . La proteína recombinante fue purificada en condiciones desnaturalizantes a partir de los cuerpos de inclusión, siguiendo el manual de columnas de nickel NI-TED2000 (Macherey-Nagel GMBH, Düren, Alemania). Para ello se resuspendieron los sedimentos recuperados a partir de 200 mL de cultivo en 10 mL tampón LEW pH 8 proporcionado por el kit y diluida a concentración 1x. La lisis celular se indujo por sonicación en un Sonifer 250/450 (Branson Ultrasonics, CT, EE.UU.), aplicando 7 pulsos de 20 s a 40% de la capacidad, alternados con 20s de incubación en hielo. Mediante centrifugación durante 30 minutos a $10.000 \times g$ y 4°C los cuerpos de inclusión se separaron de la fase soluble, la cual se descartó. El sedimento se resuspendió en otros 10 mL de tampón LEW para lavarlo y recuperarlo mediante centrifugación. Los cuerpos de inclusión se solubilizaron por adición de 4 mL tampón DS pH 8 del kit a 1x, suplementado con urea 8 M, e se incubaron durante 1 hora en hielo con agitación. Luego se centrifugó durante 30 min a $10.000 \times g$ y 20°C y el sobrenadante se pasó por filtros con poros de $0,45 \mu\text{m}$, con el propósito de eliminar contaminantes que pueden obstruir la columna de níquel. La columna se equilibró con 4 mL del tampón DS y después se cargó el extracto con las proteínas solubilizadas sobre la columna, dejándolo entrar por gravedad. Se realizaron 3 pasos de lavados, cada vez añadiendo 4 mL de tampón DS a la columna y la proteína recombinante se eluyó con 5 mL del tampón EB pH 8, que contiene imidazole 250 mM, con adición de urea 8 M.

La presencia de la proteína en las fracciones eluídas se visualizó mediante SDS-PAGE y tinción Coomassie como se describe en la sección 3.10 y la proteína recombinante fue extensivamente dializada en el tampón de extracción (HEPES 50 mM pH 7.8, KCl 50 mM, EDTA 5 mM, glicerol 10% (v/v), 2-mercaptoethanol 10 mM), para eliminar todo el imidazol y la urea. La concentración de proteínas se determinó por el método de Bradford (1976) y las proteínas purificadas se almacenaron a -80°C .

3.10 Análisis por SDS-PAGE y western blot

Para el análisis de las proteínas mediante electroforesis en SDS-PAGE y Western blot se le añadió tampón Laemmli (1970) al extracto, suplementado con 2-mercaptoethanol 0,5 mM y se incubó durante 5 minutos a 100°C para asegurar la desnaturalización de las proteínas. Se centrifugó 10 min a 18.660 x g para recoger el sobrenadante y las proteínas se separaron mediante electroforesis desnaturalizante discontinua, con un gel separador de poliacrilamida al 12,5 % (p/v) y un gel concentrador al 5 % (p/v). La electroforesis se llevó a cabo a una intensidad constante de 15 mA por gel hasta que el frente, visualizado por azul de bromofenol, había entrado en el gel separador. Entonces la intensidad se subió a 30 mA por gel durante 30 a 45 min, hasta observar una buena separación de las proteínas marcadoras preteñidas.

La visualización de la totalidad de proteínas separadas se llevó a cabo mediante una tinción durante 15 min con una disolución de azul de Coomassie R-250 al 2,75 g/L en ácido acético 10% (v/v) y metanol 40 % (v/v), seguido por varios pasos con agua destilada caliente hasta diferenciar bien las bandas. Para la tinción de plata se utilizó un kit comercial "Silver stain Plus" (Bio-Rad, Hercules, CA, EE.UU.) siguiendo las recomendaciones del manual.

Para la inmunodetección de la ASPG de pino (Western blot) las proteínas separadas por su masa mediante la electroforesis en presencia de SDS se transferían a una membrana de nitrocelulosa (Whatman GmbH, Dassel, Alemania) mediante la aplicación de un campo eléctrico constante de 25 V durante 30 min en un sistema semi-seco Trans-Blot SD Cell (Bio-Rad, Hercules, CA, EE.UU.). Después de la transferencia se procedió a la saturación de la membrana con polvo de leche al 5 % (p/v) en un tampón TBS a pH 7,4 con Tween-20 al 0,05 % (v/v) durante 1 hora. La presencia de ASPG se detectó por inmunoblot como descrito (Cañas *et al.*, 2007) usando el suero de conejo con anticuerpos policlonales, producido frente a la proteína recombinante de *P. sylvestris* y purificado como descrito (Sakakibara *et al.*,

1991) sobre la proteína recombinante de *P. pinaster*. Los anticuerpos primarios frente a la asparraginas se diluyeron 500 veces en el mismo tampón TBS con Tween-20 y polvo de leche y se incubaron una noche sobre la membrana a temperatura ambiente y con agitación suave. La mañana siguiente se aplicaron tres lavados de 10 min con tampón TBS y Tween-20 antes de añadir el anticuerpo secundario conjugado con peroxidasa contra suero de conejo. La membrana fue incubada durante 2 horas y luego se emplearon otros tres pasos de lavado. Los complejos inmunológicos se visualizaron mediante ensayo peroxidasa empleando un kit comercial "SuperSignal West Pico Chemiluminescent Substrate" (Thermo Fisher Scientific Inc., Waltham, MA EE.UU.). Las proteínas recombinantes que presentaron un tag HA a su extremo carboxilo se revelaron de la misma manera, sustituyendo el anticuerpo primario por uno elaborado contra la proteína HA.

3.11 Ensayo de actividad asparraginas/isoaspartilodipeptidasa

La actividad asparraginas se determinó midiendo la tasa de producción de amonio por la hidrólisis de L-Asn. Alícuotas de 1 μ L de la proteína se incubaron durante 30 min a 37°C en un volumen total de 15 μ L de tampón (HEPES 50 mM, pH 7.5, KCl 50 mM, asparragina 20 mM). La actividad se paró mediante incubación durante 2 min a 100°C seguida por centrifugación a 12.400 x g durante 2 min. Una alícuota de 12 μ L del sobrenadante se pasó a otro tubo para la cuantificación de amonio mediante un método basado en la reacción de fenol-hipoclorito (Loan *et al.*, 2013). Para ello a cada alícuota se añadió un volumen de 120 μ L de los siguientes reactivos en el siguiente orden: fenolato sódico 0,33M pH 13, nitroprusiato sódico a 0,01% (p/v) y lejía a 1,6 % (v/v). Después de una incubación de 30 min a temperatura ambiente se determinó la concentración de amonio por ensayo colorimétrico a 635 nm, usando una curva con diferentes concentraciones de sulfato de amonio como estándar.

La actividad isoaspartato dipeptidasa se determinó mediante la cuantificación de la liberación de aspartato, de acuerdo al ensayo descrito por Möllering (1985). Éste se basa en una primera reacción de transaminación entre aspartato y α -cetoglutarato, con la formación de oxalacetato por glutamato oxalacetato transaminasa (GOT, EC 2.6.1.1). En una segunda reacción se produce la reducción del oxalacetato a malato, catalizada por malato deshidrogenasa (MDH, EC 1.1.1.37), con la oxidación de NADH. La disminución de NADH es cuantificado mediante lectura de la absorción a 340 nm y es proporcional a la concentración de aspartato en la muestra. Las reacciones se llevaron a cabo en microplacas en un volumen total de 180 μ L con HEPES 50 mM pH 7,8, KCl 50 mM, α -cetoglutarato 3,5 mM, NADH 0,6 mM y 5 U de las enzimas GOT y MDH. Después de añadir el sustrato la mezcla se incubó durante 10 min a 25°C hasta observar la estabilización de la absorbancia a 340 nm. Luego un volumen adecuado de proteína recombinante fue añadido y la oxidación de NADH se monitorizó durante 40 min en un "PowerWave HT Microplate Spectrophotometer" (Biotek Instruments, VT, EE.UU.).

3.12 Caracterización de las proteínas recombinantes purificadas

Para la determinación de los valores K_m y V_{max} las concentraciones de los sustratos presentes en la reacción abarcaban desde 0,2 hasta 10 veces el valor de K_m . Los datos experimentales se ajustaron a una ecuación de Michaelis-Menten y el comportamiento cinético se calculó mediante el método de Hanes-Woolf ($r^2 > 0,99$).

La temperatura óptima de la proteína recombinante se ensayó desarrollando la actividad en un Biometra Tgradient thermocycler (Biometra GmbH, Göttingen, Alemania) a intervalos de 5°C en un gradiente de temperatura entre 5°C y 60°C y el producto de la reacción se cuantificó tal y como se ha descrito en la

sección 3.11. La energía de activación se calculó mediante la ecuación de Arrhenius sobre la tasa de actividad medida a las diferentes temperaturas.

Para los ensayos a diferente pH se prepararon los siguientes tampones a 50 mM: tampón fosfato pH 2, tampón acetato pH 4,5, MES pH 6, MOPS pH 7, HEPES pH 8, tampón borato pH 9. La reacción se desarrolló en las mismas condiciones como se ha descrito en la sección 3.11, pero tamponando el medio de reacción a distintos pH con las soluciones tamponadoras anteriores.

El efecto de iones monovalentes sobre la actividad asparraginasa se estimó empleando el medio de reacción ya descrito en el que se sustituyó el KCl por la sal cloruro de cada ion correspondiente a una concentración final de 50 mM. Se incluyó también una medida de actividad en una mezcla de reacción sin ninguna sal.

La estabilidad térmica se evaluó por incubación de alícuotas de la enzima purificada a distintas temperaturas (40, 50, 60 y 70°C) en el tampón de extracción (HEPES 50 mM pH 7,8, KCl 50 mM, EDTA 5 mM) en ausencia del sustrato L-Asn. Se retiraron alícuotas a distintos tiempos, las cuales se guardaron en hielo hasta la determinación de la actividad residual mediante el método descrito. El efecto del ion potasio sobre la estabilidad se determinó de manera similar, pero incubando alícuotas de la enzima durante más tiempo (3 días consecutivos) a menor temperatura (22°C y 37°C) y la actividad residual fue comparada entre las muestras que se incubaron en presencia y ausencia de KCl 50 mM.

Todos los valores se representan como media de tres réplicas experimentales \pm desviación estándar y en todos los experimentos se incluyeron reacciones control sin enzima y sin sustrato.

Los reactivos y enzimas se compraron de Sigma-Aldrich Co. LLC. (Saint Louis, MO, EE.UU.) con la excepción de β -L-Asp-L-Ala que se obtuvo de Bachem (Torrance, CA, EE.UU.) y L-Asn de Duchefa biochemie (Haarlem, The Netherlands).

3.13 Espectrometría MALDI-TOF

La proteína PpASPG1 purificada por métodos convencionales se analizó mediante espectrometría de masas MALDI-TOF (4700 Proteomics Analyzer mass spectrometer, ABSCIEX), en modo linear con el objetivo de determinar el peso molecular de las subunidades. Además, la muestra se fraccionó electroforéticamente por SDS-PAGE y las subunidades se visualizaron por tinción Coomassie. Se picaron puntos de las bandas correspondientes a ambas subunidades en el gel de acrilamida y se digirieron con tripsina para el análisis de los péptidos por MALDI-TOF. Para ello se usó el equipo en modo de reflector y de algunos de los péptidos de la ASPG identificados se confirmó la secuencia por fragmentación mediante espectrometría en tándem MALDI TOF-TOF. Estos análisis se realizaron en colaboración con la unidad de proteómica SCBI de la universidad de Málaga.

3.14 Extracción de DNA y aislamiento de la región promotora *PpASPG1*

La extracción de DNA genómico se llevó a cabo a partir de cotiledones de plántulas de *P. pinaster* tal y como se ha descrito previamente (Díaz-Moreno 2010; Villalobos et al. 2012), basado en un protocolo de Dellaporta *et al.* (1983), con la excepción de un paso adicional de extracción con cloroformo:alcohol isoamílico 24:1 después de la incubación con RNAsa. La calidad del DNA se comprobó en gel de agarosa 0,8 % (p/v). Luego se aplicó el sistema comercial "Genome Walker" (Clontech Laboratories Inc, CA, EE.UU.) para aislar la región promotora del gen *PpASPG1* de *Pinus pinaster*. Para la construcción de 6 genotecas genómicas se digirió por separado con cada una de las enzimas de restricción empleadas (*EcoRV*, *DraI*, *HpaI*, *StuI*, *PvuII* y *ScaI*) 5 µg de DNA. Todas las enzimas utilizadas generan extremos romos. Se usaron 80 U de las enzimas en un volumen de 100 µL y las reacciones se

incubaron una noche a 37°C. Después se aisló el DNA digerido siguiendo las instrucciones del manual. Se unieron los adaptadores a los extremos mediante una reacción de ligación que incluyó: 4 µL del producto de digestión con 3 U de la T4 DNA ligasa, su tampón a concentración 1x y los adaptadores a 5,9 µM. La mezcla se incubó durante 12 horas a 16°C y luego las genotecas llamadas DL1 a DL6 se almacenaron a -20°C.

Para evitar diseñar los cebadores en una región interrumpida por un intrón, primero se identificó la posición probable del primer intrón en la secuencia genómica del gen de la ASPG de *P. pinaster*. Para ello se buscaron genes ortólogos en los genomas de *Arabidopsis thaliana* y *Populus trichocarpa* y se alinearon con la secuencia completa del cDNA de *PpASPG1* usando el programa CLUSTAL Omega (Sievers *et al.*, 2011). Una vez identificada la posición del primer intrón que estaba conservada en las ASPG de las otras especies, se diseñó una pareja de cebadores AP-1 y AP-2 flanqueando esta posición para amplificar un fragmento de DNA genómico que incluía el primer intrón. La reacción de PCR se hizo usando 250 ng del DNA genómico extraído de cotiledones como molde en un volumen de 50 µL conteniendo dNTPs a 0,2 mM, los cebadores externos a 0,2 µM y 2,5 U de "Advantage genomic LA polimerase" (Clontech Laboratories Inc, CA, EE.UU.) en el tampón proporcionado por la casa comercial diluida a concentración 1x. El programa de PCR tenía los siguientes pasos: 1 min a 94°C de desnaturalización inicial, 30 ciclos de 30 s a 94°C, 30s a 57°C y 3 min a 68°C, seguido por un paso final de 10 min a 72°C. El fragmento amplificado se purificó, se ligó en el vector pBluescript SK+ y se transformaron células de *E. coli* para la propagación del plásmido y secuenciación del inserto, tal y como se ha descrito previamente en la sección 3.4.

Sobre la secuencia conocida del primer exón e intrón de la *PpASPG1* se diseñó un juego de 2 oligonucleótidos anidados para amplificar la región promotora, el cebador externo posicionado sobre la zona de transición exón-intrón (AP-3) para aumentar la especificidad para el gen funcional de ASPG, y el segundo cebador interno más hacia 5'UTR (AP-4). La región promotora se consiguió

aislar por dos reacciones seguidas de PCR sobre las 6 genotecas, usando los cebadores específicos para el gen de ASPG en combinación con los proporcionados por el kit que hibridan a la secuencia de los adaptadores. En una primera PCR se usó 1 µL de las genotecas como molde en reacciones de PCR con las mismas condiciones como especificadas usando 2,5 U de la DNA polimerasa "Advantage genomic LA polimerase". El programa de PCR era 1 min a 94°C, 7 ciclos de 25 s a 94°C y 5 min a 72°C, 32 ciclos de 25s a 94°C y 5 min a 68°C y un paso final de 7 min a 72°C.

Los productos de la primera reacción de PCR se diluyeron 50 veces con agua estéril y se usó 10 µL de la dilución como molde en la segunda PCR con los cebadores internos y un programa de 1 min a 94°C, 5 ciclos de 25 s a 94°C y 5 min a 72°C, 20 ciclos de 25s a 94°C y 5 min a 68°C y un paso final de 7 min a 72°C.

Los fragmentos se clonaron en el vector pBluescript SK+ y se secuenciaron. La secuencia genómica se obtuvo a partir de la secuenciación de tres fragmentos amplificados por el procedimiento descrito en eventos independientes de PCR.

3.15 Análisis *in silico* de la región promotora *PpASPG1*

Las regiones promotoras del gen *PpASPG1* se secuenciaron en un secuenciador Beckman Coulter CEQ8000 y sobre la secuencia obtenida se buscaron posibles elementos reguladoras *cis* mediante las siguientes aplicaciones disponibles en la red:

PlantCARE

(<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>)(Lescot *et al.*, 2002) y PLACE (<https://sogo.dna.affrc.go.jp/cgi-bin/sogo.cgi?lang=en>) (Higo *et al.*, 1999).

3.16 Análisis funcional de la región promotora *PpASPG1*

Para ensayar la funcionalidad de la región promotora aislada se optó por el vector binario pBI121 que permite comprobar la expresión transitoria del gen delator dirigido por el promotor en hojas de *N. benthamiana* por transformación con *Agrobacterium tumefaciens*. El procedimiento seguido se resume en el lado derecho de la figura 3.4. Dado el bajo nivel de expresión del gen *PpASPG1* en acículas de pino, no esperábamos obtener grandes cantidades de proteína delatora en un sistema heterólogo bajo el promotor de pino. Por eso se decidió sustituir el gen delator en el vector por otro que permitiera mayor sensibilidad en los ensayos. Para ello, el vector pBI121 se digirió con las enzimas de restricción *BamHI* y *SacI* Fast Digest (Thermo Fisher Scientific Inc, Waltham, MA, EE.UU.), eliminando la región codificante para GUS y la secuencia codificante para luciferasa fue introducido para la construcción del vector control positivo con el gen codificante para luciferasa bajo el control del promotor CaMV 35S.

Con el propósito de fusionar la secuencia promotora del gen *PpASPG1* con la del gen delator luciferasa para los ensayos de expresión transitoria en hojas de *N. benthamiana*, se diseñaron oligonucleótidos que amplifican la secuencia aislada del 5' hasta la posición del inicio de la transcripción. Los cebadores AP-5 y AP-6 en combinación con AP-8 producen las dos versiones completas de promotor aislado e incorporan los sitios de restricción *HindIII* y *BamHI* en los extremos 5' y 3' respectivamente. Como molde se usaron los plásmidos pBSK que contenían los fragmentos promotores previamente aislados. Mediante una reacción de PCR con cebadores AP-7 y AP-8 se generó una versión más corta del promotor, abarcando solamente la secuencia del promotor proximal de 476 nt.

Los productos de PCR se digirieron con las enzimas de restricción *HindIII* y *BamHI* y se ligaron en el vector pBI121 igualmente digerido, sustituyendo la secuencia promotora CaMV 35S. De tal manera la región promotora del gen *PpASPG1* queda posicionada a 5' de la secuencia codificante para luciferasa y

dirige su expresión. Todas las construcciones se emplearon para transformar la cepa DH5a para propagación del plásmido y verificación de la construcción por secuenciación.

La transformación de *A. tumefaciens* y la expresión transitoria del gen de luciferasa en hojas de *N. benthamiana*, dirigido por los diferentes promotores, se hizo de la misma manera como se describe en la sección 3.6 y se representa en la figura 3.4. Para el análisis de los niveles de expresión, las zonas de las hojas infiltradas se recolectaron tras tres días después de la agroinfiltración y la actividad luciferasa se cuantificó empleando el sistema comercial “Luciferase Assay System” (Promega Corp, Madison, WI, EE.UU.) como indica el manual. La luminiscencia fue registrada en un luminómetro Biotek FLx800 (Biotek Instruments, VT, EE.UU.). Como control se determinó la actividad luciferasa presente en los cultivos de *A. tumefaciens* previamente a la agroinfección.

3.17 Aislamiento de la región promotora *PpASPG1* en pinos de diferentes procedencias

Se recolectaron acículas de plantas de *P. pinaster* de 11 poblaciones distribuidas entre España, Francia y Marruecos (Figura 3.2) y se extrajo el DNA genómico empleando un sistema comercial “NucleoSpin Plant II” (Macherey-Nagel GmbH, Düren, Alemania) y siguiendo las instrucciones del manual. Las acículas se homogeneizaron por maceración en nitrógeno líquido y 100 mg de tejido pulverizado fue mezclado con 600 µL de tampón de lisis PL1 (basado en un protocolo de CTAB) proporcionado por el sistema comercial. Se añadieron 150 µg de RNasa del kit y tras 20 min de incubación a 65°C la mezcla se depositó sobre unos filtros para aclarar el lisado mediante centrifugación. Posteriormente el filtrado se cargó sobre las columnas de silica y tras hacer pasar tres veces un tampón de lavado proporcionado por el kit, se eluyó el DNA purificado en 50 µL de tampón Tris/HCl 5 mM, pH 8,5.

Se realizaron reacciones de PCR con el DNA extraído como molde, usando los cebadores AP-9 diseñado sobre la región promotora proximal en dirección sentido y AP-10 en sentido contrario sobre el primer exón de la región codificante del gen *PpASPG1*. Las reacciones de PCR se hicieron en un volumen de 25 µL con 400 ng de DNA como molde y a las siguientes condiciones: 2 min a 98°C de desnaturalización inicial seguida por 30 ciclos de 10 s a 98°C, 30s a 57°C y 1 min a 72°C, con un paso final de 10 min a 72°C. Los fragmentos amplificados se clonaron en el vector pBSK para su secuenciación.

3.18 Reacciones de PCR cuantitativo en tiempo real

Para el proceso de qRT-PCR se optó por el método de dos pasos: en el primer paso se sintetizó cDNA y en el segundo paso se desarrollaron las reacciones de PCR. El RNA total se aisló de los tejidos muestreados y se sintetizó cDNA como se describió en la sección 3.3.

Se diseñaron parejas de oligonucleótidos sobre las secuencias de los diferentes genes de asparraginas que se habían clonado, posicionando uno de ellos en la zona codificante en dirección 3' y el otro en la región no codificante 3'UTR del cDNA en dirección hacia 5', aumentando así la especificidad de amplificación del fragmento para el gen en cuestión. El tamaño del amplicón siempre se limitó al rango de 120 a 250 nucleótidos para minimizar el efecto de la longitud del producto sobre la eficiencia de la reacción. Además, se prestó atención en que los cebadores cumplieron los requisitos de tener un % GC alrededor de 50 %, que la temperatura de hibridación de cada pareja fuera igual con una diferencia máxima de 1°C, y que no se formaran horquillas o dímeros entre ellos o consigo mismo. Se emplearon las parejas de cebadores A1-13 con A1-14 para *PpASPG1*, A2-5 con A2-2 para *PpASPG2* y A3-5 con A3-6 para *PpASPG3*. Para analizar los niveles de transcritos de los genes δ -OAT y *GS1b* se diseñaron los oligonucleótidos OA-1 con OA-2 y GS-1 con GS-2 respectivamente, sobre la secuencia disponible en la base de datos

SustainPineDB v.3.0, unigenes 5775 y 10503 respectivamente (http://www.scbi.uma.es/sustainpinedb/home_page).

La amplificación cuantitativa qRT-PCR se realizó en un iCycler iQ System (Bio-Rad, Hercules, CA, EE.UU.), utilizando un procedimiento basado en cuantificación de la fluorescencia de SYBR-Green y empleando la mezcla de reacción "iQ SYBR Green Supermix", también suministrada por Bio-Rad. Las reacciones se desarrollaron en un volumen de 25 µL con 4 µL de cDNA a una concentración adecuada (2,5 ng/µL, 5 ng/µL o 12,5 ng/µL), los dos cebadores específicos a 0,4 µM y el reactivo comercial a 1x que contiene Tris-HCl 20 mM pH 8,4, KCl 50 mM, dNTPs a 0,2 mM cada uno, iTaq DNA polimerasa 25 U/mL, MgCl₂ 3 mM, fluoresceína 10 nM, SYBR Green I y estabilizadores. Para cada muestra, las reacciones se realizaron por triplicado y se incluyó un control con RNA sin retrotranscribir. En cada ensayo y para cada pareja de cebadores se incluyó, además, un control sin molde. El programa de PCR consistió en una desnaturalización inicial de 3 min a 95°C, seguida por 50 ciclos de 15 s a 95°C, 20 s a 60°C y 30s a 72°C para el gen *PpASPG1* y todos los genes de referencia. Los ensayos de los otros genes se realizaron a las mismas condiciones con excepción del paso de hibridación que se hizo a una temperatura de 59 °C para *PpASPG2* y *PpASPG3* y a 58°C para *δ-OAT* y *GS1b*. Después del ciclo final se añadió un análisis de disociación en un rango de temperatura de 55–95°C con incrementos de 0.5°C para verificar la especificidad de las reacciones.

Los datos de fluorescencia se obtuvieron por el módulo de análisis de Bio-Rad iQ5 en modo crudo (raw data) y se analizaron mediante la aplicación LinRegPCR software versión 11.0 (Ruijter *et al.*, 2009) que determina la fase exponencial para cada reacción independiente, además de calcular los parámetros de eficiencia de la PCR (E) y el ciclo umbral (Ct). La expresión relativa del gen (R0) se calculó mediante la fórmula

$$R0 = 1/(1+E)^{Ct}$$

donde E es la eficiencia media de las eficiencias de todas las reacciones de un mismo gen analizado en la misma placa y Ct es el ciclo umbral para cada réplica independiente. Luego se calculó R0 para cada muestra como la media de los tres R0 de las tres réplicas técnicas.

Los valores de expresión relativa se normalizaron con los genes de referencia actina, ubiquitina, el factor de elongación EF1- α , y las subunidades L34 y S27 del complejo ribosomal de la siguiente manera: se calculó la razón entre la expresión relativa del gen candidato (R0) y la media geométrica entre los valores R0 de los genes de referencia más estables NF (factor de normalización). La desviación estándar se determinó mediante la fórmula

$$SD_{R0/NF} = R0/NF \times \sqrt{[(SD\ NF/ NF)^2 + (SD\ R0/R0)^2]}.$$

Los genes más estables a incluir para la normalización se identificaron mediante el algoritmo "geNorm" (Vandesompele *et al.*, 2002) y se aplicó una estadística de contraste *t* de Student con un nivel de confianza del 95 %.

El estadístico (*t*) y los grados de libertad (*v*) se calcularon de la siguiente manera:

$$t = (M_a - M_b) / \sqrt{(SD_a^2/n_a + SD_b^2/n_b)}$$

$$v = (SD_a^2/n_a + SD_b^2/n_b)^2 / [(SD_a^2/n_a)^2 / n_a - 1 + (SD_b^2/n_b)^2 / n_b - 1]$$

donde a y b son las muestras comparadas, M es la media de los valores de expresión (R0/NF), SD es la desviación estándar y n es el número de réplicas (n=3).

Para los resultados que se presentan como el cambio en nivel de mRNA relativo entre las muestras de los tratamientos con respecto a las plantas control, se dividió la media de expresión relativa en las plantas sometidas a los tratamientos (RT) por la media de expresión relativa en las plantas control (RC) y se determinó la desviación estándar mediante la fórmula

$$SD_{RT/RC} = RT/RC \times \sqrt{[(SD\ RT/RT)^2 + (SD\ RC/RC)^2]}$$

Luego se aplicó Log en base de 2 sobre la razón RT/RC y la desviación estándar se propagó mediante la fórmula

$$SD \log_2 (RT/RC) = SD (RT/RC) / [(RT/RC) \cdot \ln 2]$$

4. RESULTADOS

4.1 Identificación de asparraginasas en el transcriptoma de *Pinus pinaster* Aiton.

4.1.1 Aislamiento de secuencias de cDNA de longitud completa

Entre los diversos genomas de plantas que han sido secuenciados, las bases de datos incluyen muy pocas especies gimnospermas. Una búsqueda de secuencias codificantes para asparraginasas en los genomas y EST publicados de coníferas produce pocos resultados de secuencias completas. Como punto de partida para este trabajo se identificaron secuencias de asparraginasa presentes en el transcriptoma de la especie *Pinus pinaster* Aiton y se obtuvieron los correspondientes cDNA completos. A partir de una secuencia EST previamente aislada por nuestro grupo de trabajo (Villalobos *et al.*, 2012) se empleó una estrategia RLM-RACE para amplificar el cDNA de longitud completa de una asparraginasa que denominamos *PpASPG1*. La secuencia del fragmento aislado se compone de una región 5'UTR de 129 nucleótidos, un ORF de 1128 pares de bases (incluyendo el codón STOP) y una región 3'UTR de 139 nucleótidos, seguidos por una secuencia poli-A. Usamos esta secuencia para buscar posibles genes de asparraginasa mediante un BLAST en el transcriptoma de *P. pinaster* que fue publicado por Canales *et al.* (2014). Entre los resultados encontramos un unigen de secuencia idéntica al cDNA de *PpASPG1* (sp_v3.0_unigene4029) además de dos más que codifican para asparraginasas, unigenes 15077 y 36994, los cuales denominamos *PpASPG2* y *PpASPG3* respectivamente. Otros resultados correspondieron a fragmentos de longitud parcial o de inferior similitud (valor $e > 10^{-50}$) y más parecidos a aspartilglucosaminidasas o treonina aspartasas. Las secuencias de longitud completa para los genes *PpASPG2* y *PpASPG3* se amplificaron sobre cDNA sintetizado a partir de RNA extraído de plántulas jóvenes. Ambas secuencias

contienen un ORF de 921 nucleótidos que codifica para una proteína de 306 aminoácidos, notablemente más corta que PpASPG1 con 375 residuos. Las regiones 5'UTR de *PpASPG2* y *PpASPG3* presentan un 67,50 % de identidad entre ellos y por término medio un 42,79 % con *PpASPG1*. En 3'UTR el % de identidad entre los tres genes se sitúa entre 45,38 % y 55,13 %. Las secuencias primarias de PpASPG2 y PpASPG3 se parecen más entre ellas que a PpASPG1 (Figura 4.2). El % de identidad de aminoácidos de 83,28 % (Figura 4.3) y una similitud de 90,16 % indica que PpASPG2 y PpASPG3 son proteínas diferentes. Concluimos que en la especie *Pinus pinaster* existen, al menos, tres genes que codifican proteínas diferentes con similitud a las asparraginasas. En este trabajo se presentan algunos aspectos sobre sus características y regulación con el objetivo de determinar su posible función durante el desarrollo de las plántulas.

4.1.2 Estructura primaria de las ASPG de coníferas

A partir de los cDNAs aislados se dedujeron las estructuras primarias de las tres asparraginasas y en la figura 4.1 se presenta un alineamiento que incluye también asparraginasas de otras coníferas cuya secuencia está completa (*P. sylvestris* y *P. taeda*), además de las de algunas especies angiospermas que han sido caracterizadas. La comparación de los aminoácidos permite hacer algunas suposiciones sobre las características de las asparraginasas de pino, las cuales se pretenden comprobar experimentalmente más adelante en este trabajo. La estructura de la ASPG dependiente de potasio de *Phaseolus vulgaris* (PvAspG1) fue publicada recientemente (Bejger *et al.*, 2014) y en base a las conclusiones de los autores se han marcado sobre el alineamiento la posición de los residuos involucrados en la catálisis. Las otras asparraginasas incluidas como referencias en las figuras han sido clasificadas como dependientes o independientes del ion potasio en diversos trabajos (Borek *et al.*, 2004; Bruneau *et al.*, 2006; Credali *et al.*, 2011; Gabriel *et al.*, 2012).







-  N-terminal nucleófilo
-  Anclaje del producto
-  Especificidad de sustrato
-  Estabilización por Na⁺
-  Activación por K⁺
-  Interruptor catalítico

Figura 4.1 Alineamiento de las tres asparraginasas de *Pinus pinaster* con las enzimas de otras coníferas y de especies angiospermas caracterizadas (Borek *et al.*, 2004; Credali *et al.*, 2011; Gabriel *et al.*, 2012; Bejger *et al.*, 2014). La secuencia de aminoácidos de las tres asparraginasas de *P. pinaster* identificadas se alinearon mediante la aplicación Clustal Omega (Sievers *et al.*, 2011) y se incluyeron algunas asparraginasas de angiospermas que han sido experimentalmente caracterizadas como dependientes del ion potasio (At3g16150 de *A. thaliana*, PvAspG1 de *P. vulgaris*, LjNSE1 de *L. japonicus*) e independientes del ion (At5g08100 de *A. thaliana* y LIA de *L. luteus*) y de las coníferas *P. taeda* (Ptaeda) y *P. sylvestris* (Psylv), cuyas secuencias están completas en las bases de datos pero que no han sido caracterizadas. Las regiones con residuos identificados por Bejger *et al.* (2014) como importantes para la actividad catalítica están marcadas con los colores según especifica la leyenda. La treonina marcada en color gris es el residuo nucleófilo y marca también el sitio de procesamiento que divide el precursor en dos subunidades de tal manera que la treonina se localiza en el extremo N-terminal de la subunidad β .

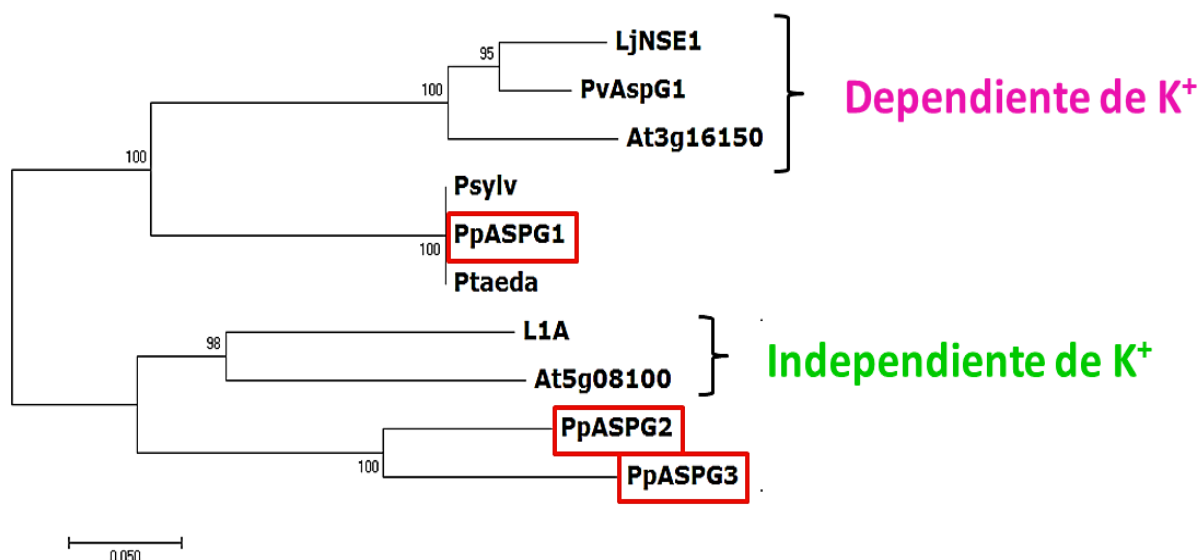


Figura 4.2 Clasificación de las asparraginasas de coníferas y especies angiospermas en base a su estructura primaria. Se elaboró un árbol filogenético usando el método "neighbour joining" de la aplicación MEGA6 (Tamura *et al.*, 2013) sobre las secuencias alineadas de la figura 4.1. Se representa el árbol consenso de 500 repeticiones y el porcentaje de réplicas obtenidas se indica al lado de cada rama. Las distancias evolutivas se calcularon por el método de Poisson (Zuckerkandl and Pauling, 1965) y se reflejan en la longitud de las ramas según la escala mostrada debajo del árbol.

Los aminoácidos relacionados con el anclaje del producto en el sitio activo y del interruptor catalítico están perfectamente conservados entre todas las secuencias, pero los residuos determinantes de la especificidad de sustrato (marcados en azul) presentan gran variabilidad, aunque con un mayor grado de similitud entre las enzimas del mismo subgrupo. La comparación de estos residuos sugiere que PpASPG1 es una enzima ortóloga a las asparraginasas dependientes de potasio en angiospermas mientras que PpASPG2 y PpASPG3 son más parecidas a las asparraginasas independiente de potasio. La dependencia de este ion según Bejger *et al.* (2014) está vinculada con la secuencia VMDKSPHS (marcada en amarillo) de la cual el último residuo no se conserva en PpASPG1, ni en los ortólogos de *P. sylvestris* y *P. taeda*. En la región que forma un bucle que interactúa con un ion sodio, importante para la estabilidad de la enzima (marcada en color verde), también se aprecia la división de las secuencias en dos subgrupos. Dado la divergencia en secuencia entre las asparraginasas de coníferas y sus ortólogas de angiospermas, podría haber diferencias en la preferencia de sustrato. Por otro lado, los cambios en el motivo VMDKSPHS observados en la secuencia de PpASPG1 con respecto a las ASPG dependientes de potasio de especies angiospermas sugiere que sería de interés estudiar el efecto del ion potasio sobre la actividad enzimática en esta enzima de pino.

La treonina que en angiospermas está identificada como el residuo nucleófilo con actividad catalítica está marcada en gris en la figura 4.1 y permite identificar el sitio de procesamiento que divide el precursor en dos subunidades. En el apartado 4.3.1 de este trabajo se demuestra la conservación de la posición de la autocatálisis con las proteínas homólogas de especies angiospermas. Aunque la secuencia primaria de los precursores está bien conservada, destaca una diferencia notable en aminoácidos y en longitud en la secuencia que precede al sitio de autoproteólisis. Esta región es excepcionalmente extensa en el caso de PpASPG1 y las proteínas ortólogas de otras coníferas, superando con un total de 375 residuos a sus homólogos con 50 a 70 aminoácidos en longitud. En las secciones 4.3 y 4.4 se analizan los

efectos que modificaciones en esta RV tienen en el proceso de la maduración y/o las características bioquímicas de PpASPG1.

% identidad entre las ASPG de pino

	PpASPG2	PpASPG3
PpASPG1	64,47	63,16
PpASPG2	-	83,28

% identidad con ASPG de angiospermas

	At3g16150	PvAspG1	LjNSE1	At5g08100	L1A
PpASPG1	69,14	69,54	68,62	63,02	62,93
PpASPG2	60,86	60,2	60,53	68,75	69,18
PpASPG3	57,57	57,57	58,22	67,76	67,54

% identidad

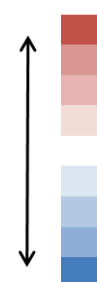


Figura 4.3 Porcentaje de identidad entre las secuencias primarias de diferentes asparraginasas. Se representa el % en identidad de aminoácidos calculados sobre el alineamiento de las secuencias de la figura 4.1. Las asparraginasas de angiospermas identificadas en color fucsia se caracterizaron como dependiente de potasio y las de color verde pertenecen al otro subgrupo.

4.1.3 Clasificación filogenética de las asparraginasas de pino

La comparación de las secuencias en el alineamiento de la figura 4.1 sugiere un mayor parecido de PpASPG1 con el subgrupo de las asparraginasas dependientes de potasio y un árbol filogenético elaborado a partir de este alineamiento confirma esta afirmación. El análisis presentado en la figura 4.2 agrupa la PpASPG1 con el subgrupo dependiente de potasio. La secuencia primaria de PpASPG1 tiene un 69,14% de identidad (Figura 4.3) con la At3g16150 que ha sido caracterizada experimentalmente como una asparraginasa dependiente del ion potasio (Bruneau *et al.*, 2006), mientras que las otras ASPG de *P. Pinaster* identificadas tienen mayor identidad con las

asparraginasas independientes de potasio. Este resultado sugiere que la separación en dos subgrupos de asparraginasas de plantas ocurrió antes de la diversificación de especies angiospermas y gimnospermas.

4.2 Actividad catalítica de las tres asparraginasas de pino

4.2.1 Expresión de las proteínas recombinantes *in planta* y comparación del proceso de maduración

Después de múltiples intentos sin éxito de purificar las asparraginasas de tejidos de pino, decidimos expresar las proteínas recombinantes en un sistema heterólogo para poder estudiarlas. Las tres isoformas de ASPG se expresaron transitoriamente en hojas de *N. benthamiana*, mediante la técnica de agroinfección, utilizando un vector pJL-TRBO para conseguir un alto nivel de expresión de las proteínas recombinantes. El esquema de la figura 4.4a representa las diferentes construcciones que se insertaron en el vector para su expresión *in planta*. Las tres proteínas se expresaron en forma de precursor y las subunidades que se producirían en el caso de que el procesamiento ocurriese en la posición descrita para las asparraginasas de angiospermas están marcadas en diferentes colores. Las construcciones incluían en el extremo 3' del marco abierto de lectura dos secuencias marcadoras para facilitar la purificación e identificación con anticuerpos de las proteínas producidas (HA y poli-His). Cinco días después de la agroinfección se extrajeron las proteínas solubles de las zonas de hojas infiltradas y se determinó la actividad asparraginasa. En la figura 4.4b se representan los valores de la actividad medida en los diferentes extractos. Los resultados indican que las tres asparraginasas identificadas en el transcriptoma del pino tienen actividad asparraginasa y por ende los precursores han sufrido un procesamiento *in planta* que produce la enzima activa. La figura 4.4c muestra el resultado del análisis por western blot de los extractos utilizando un

anticuerpo contra la secuencia marcadora HA para identificar las proteínas recombinantes. En los tres extractos, correspondientes a la expresión heteróloga de las tres asparraginasas de *P. pinaster*, se detectó una banda a la altura del peso teórico calculado para la subunidad β , lo que confirma la correcta maduración y activación de las tres enzimas *in planta*. No obstante, en el caso de las proteínas recombinantes PpASPG2 y PpASPG3 se detectó además un polipéptido con masa molecular correspondiente al precursor. Por lo tanto, las tres asparraginasas de pino producidas *in planta* en un sistema heterólogo se procesaron mediante hidrólisis en el sitio de procesamiento conservado, aunque el procesamiento fue aparentemente más eficiente en el caso de PpASPG1.

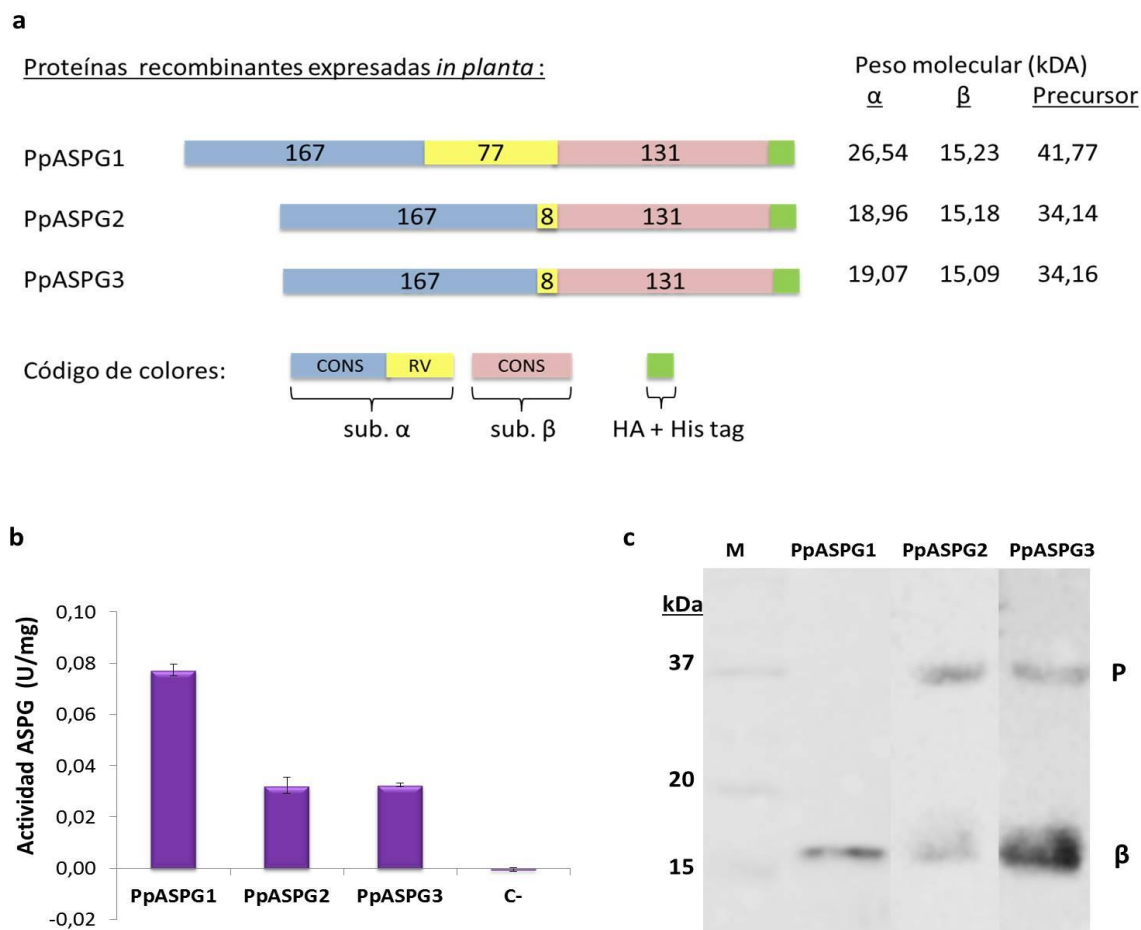


Figura 4.4 Expresión de las tres ASPG de pino *in planta* y comprobación de su actividad y proceso de maduración. Las proteínas recombinantes se expresaron transitoriamente en hojas de *Nicotiana benthamiana*. a) Representación de las diferentes formas expresadas. Se hicieron construcciones con los tags HA y poli-His en el extremo carboxilo del precursor. Las subunidades que resultan del procesamiento del precursor están marcadas en diferentes colores y el número en la subunidad indica el número de aminoácidos que forman parte de las mismas, separado en regiones conservadas con enzimas ortólogas (CONS) y región de secuencia variable (RV). La tabla a la derecha indica el peso molecular teórico de las subunidades y del precursor, en los cuales se considera el peso que suman los tags. b) Tras la extracción de las proteínas solubles de los tejidos agroinfiltrados se midió actividad asparraginasa con L-Asn 20 mM como sustrato. Como control negativo se incluyó el extracto de una hoja infiltrada con el vector pJL-TRBO sin inserto de cDNA codificante en el sitio de clonaje (C-). Se representa la actividad media de tres replicas en U por mg de proteína total \pm SD. c) Western blot de los extractos de hojas de *N. benthamiana* inoculadas con *A. tumefaciens*. Se analizó 50 μ g de proteínas totales por pocillo. Las proteínas recombinantes se detectaron mediante un anticuerpo contra el tag localizado en el extremo carboxilo y el tamaño aparente permite diferenciar entre los precursores no procesados y las subunidades β , productos del procesamiento.

4.2.2 Determinación de la preferencia de sustrato

Para poder caracterizar funcionalmente las asparraginasas de pino se procedió con la purificación de las proteínas recombinantes desde los extractos de *N. benthamiana*. Las enzimas PpASPG2 y PpASPG3 se purificaron mediante cromatografía de afinidad en columnas de Ni^{2+} , que retienen la proteína recombinante debido al tag poli-His que se añadió en sus extremos carboxilo. Después de la elución y diálisis se recuperaron 4,73% y 8,70% de la actividad inicial con factores de purificación de 7,31 y 10,50 para PpASPG2 y PpASPG3 respectivamente. No se logró obtener una purificación de PpASPG1 por este método por lo cual se expresó la proteína en su forma nativa sin tag *in planta* y se purificó siguiendo un protocolo convencional con tres pasos de purificación. Después de la concentración de la proteína en $(\text{NH}_4)_2\text{SO}_4$ al 80% (p/v), extracción con EtOH al 60% (v/v) y cromatografía de intercambio iónico el rendimiento final fue de 1,97% con un factor de purificación de 10,73.

Se compararon las actividades específicas de las preparaciones de las tres ASPG para dos sustratos diferentes, L-Asn y β -Asp-Ala, con el objetivo de evaluar las preferencias de ambos sustratos (Figura 4.5). La velocidad de reacción se mantuvo lineal durante al menos 90 minutos a 37°C y fue proporcional a la cantidad de enzima añadida para las tres isoformas. PpASPG1 presentó una mayor preferencia para L-Asn, ya que a una concentración de 25 mM de este sustrato se registró una actividad de 2,15 U/mg, mientras que a la misma concentración de β -Asp-Ala el valor registrado fue de 0,45 U/mg. Cuando se aumentó la concentración de L-Asn se observó una ligera disminución de la actividad, probablemente debido a una inhibición por exceso de sustrato. Curiosamente esto no ocurrió con el otro sustrato e incluso a una concentración de 100 mM no se llegó a saturación. PpASPG3 muestra un comportamiento contrario: se observó una velocidad de reacción mayor para β -Asp-Ala que para L-Asn, dentro de un rango de concentración entre 1 y 25 mM de sustrato. Además, presentó mayor afinidad para el primer sustrato con

el que se alcanzó la saturación entre 5 y 10 mM. Los ensayos de la proteína PpASPG2 muestran una baja actividad con ambos sustratos.

La figura 4.6a resume la actividad catalítica en una sola gráfica para comparar la especificidad de sustrato de cada ASPG. A la concentración de 25 mM en la que PpASPG1 presentó la máxima actividad para L-Asn, PpASPG2 y PpASPG3 tenían sólo un 11% y 20% del valor determinado para PpASPG1, respectivamente. Con el sustrato β -Asp-Ala la PpASPG1 y PpASPG2 mostraron un 23% y 16% de la actividad de PpASPG3. La tabla (Figura 4.6b) contiene los valores de K_m y V_{max} calculados. Se incluye también la razón V_{max}/K_m como indicador de la eficiencia de las enzimas para los sustratos. En el caso de PpASPG2 los valores registrados no permitieron calcular de forma fiable las K_m . La representación por el método de Hanes-Woolf sobre los datos de PpASPG1 con β -Asp-Ala produjo una recta de regresión con $r^2 = 0,9964$ y la ecuación asociada asigna un valor K_m de 56,73 mM. Interpretamos este dato como un valor aparente ya que no fue posible ensayar la actividad a una concentración de $10 \times K_m$. Considerando los resultados de esta caracterización concluimos que la principal función de PpASPG1 consiste en catalizar la hidrólisis de L-Asn, mientras que el papel de PpASPG3 está relacionado con la hidrólisis de dipéptidos de isoaspartilo. La diferencia en la velocidad de reacción de PpASPG2 para los dos sustratos ensayados no fue significativa según el test estadística t de Student ($P \leq 0,05$). Suponemos que la PpASPG2 es una enzima de poca eficiencia o posiblemente su principal función está relacionada con un sustrato que no fue ensayado.

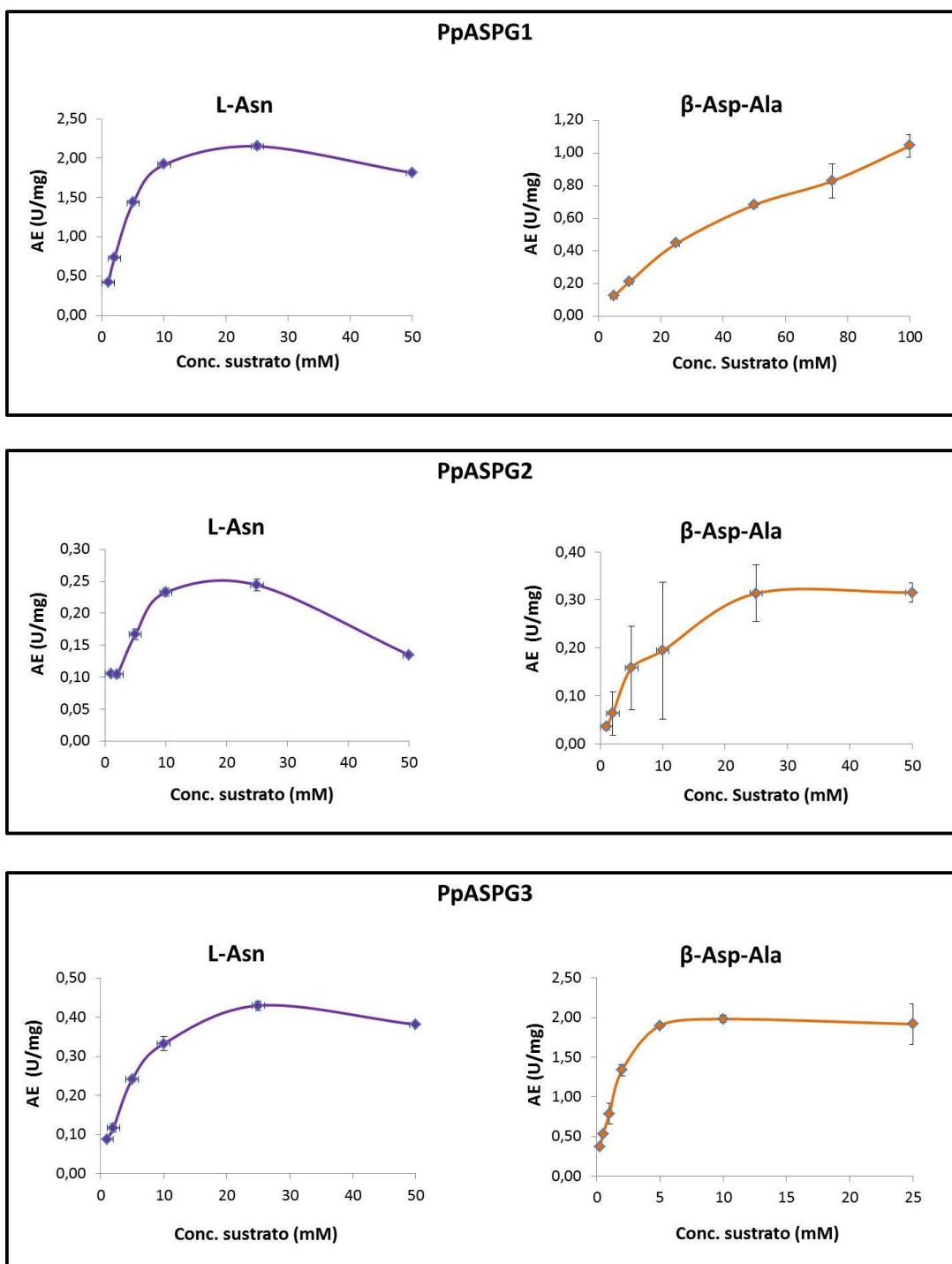
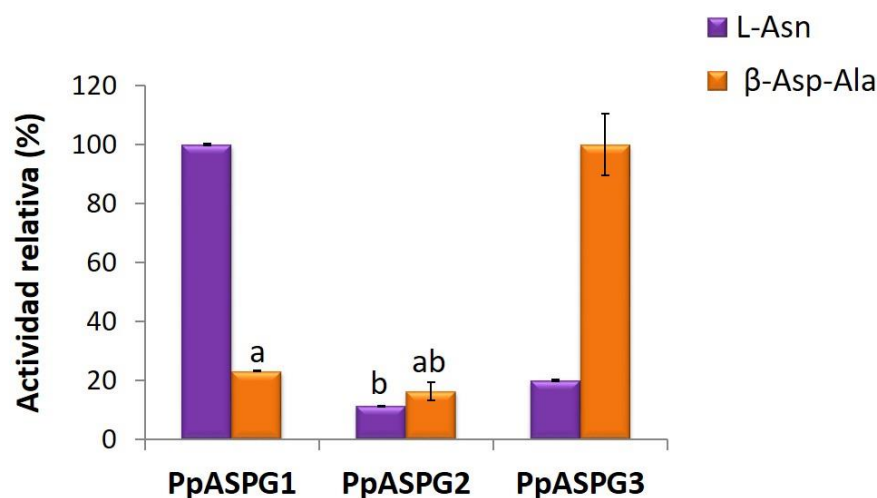


Figura 4.5 Actividad catalítica de las ASPG de pino con dos sustratos diferentes.

Las proteínas recombinantes se purificaron de hojas de *N. benthamiana* agroinfectadas para su caracterización. Se determinó la actividad específica (AE) de las tres asparaginidas purificadas para los sustratos L-Asn y β -Asp-Ala a diferentes concentraciones. Se representa la cinética de cada enzima para los dos sustratos por separado como media de tres medidas \pm SD.

a



b

	Sustrato	K _m (mM)	V _{max} (U/mg)	V _{max} /K _m (U/mg.mM)
PpASPG1	L-Asn	4,58 ± 0,03	2,59 ± 0,01	0,56
	β-Asp-Ala	56,73 ± 13,88 *	1,46 ± 0,28 *	0,03
PpASPG2	L-Asn	ND	0,27 ± 0,01	ND
	β-Asp-Ala	ND	0,37 ± 0,05	ND
PpASPG3	L-Asn	6,12 ± 0,96	0,53 ± 0,01	0,09
	β-Asp-Ala	1,04 ± 0,21	2,03 ± 0,23	1,96

Figura 4.6 Especificidad de sustrato y parámetros cinéticos para las tres ASPG de pino.

a) Comparación de la actividad específica de las tres asparagininasas. Para cada sustrato a la mayor actividad registrada se le asignó el valor de 100% y las actividades relativas de las otras enzimas para el sustrato a la misma concentración se expresan como % de ello. Los valores son medios entre tres réplicas ± SD. Se realizó una estadística *t* de Student ($P \leq 0,05$) y los valores marcados con las mismas letras no son significativamente diferentes. Los demás valores son significativos. b) De las curvas de actividad (figura 4.5) que se ajustaron a una cinética de Michaelis-Menten se calculó la K_m y V_{max} mediante una representación por el método Hanes-Woolf. La velocidad de PpASPG1 con el sustrato β-Asp-Ala no llegó a saturarse a la concentración máxima ensayada y se mencionan los valores aparentes marcados con *. La baja velocidad de PpASPG2 con ambos sustratos no permitió determinar K_m (ND) y se indica la máxima velocidad detectada durante el experimento como V_{max} .

4.3 Estructura de la PpASPG1 madura y mecanismo de activación

4.3.1 Análisis de la proteína PpASPG1 madura

El elevado número de residuos presentes en la RV de la subunidad α de las asparraginasas dependientes de potasio en coníferas nos hizo plantearnos si esta región forma parte de la proteína activada. Con el objetivo de confirmar además que el sitio de procesamiento se encuentra en el extremo amino del residuo de treonina equivalente al de las asparraginasas de angiospermas, se analizó la enzima madura mediante espectrometría de masas. Como esta técnica requiere de la proteína en una forma más pura, incluimos un cuarto paso en el protocolo de purificación de la proteína recombinante producida *in planta*, consistente en una cromatografía de filtración en gel, tras la cual la proteína recombinante PpASPG1 se había purificado con un factor de 26,60 y un rendimiento final de 0,57%.

Se realizó un análisis de espectrometría de masas MALDI-TOF en modo linear sobre la muestra pura sin digerir, con el objetivo de determinar el peso molecular de las subunidades que forman parte de la proteína madura activada. En el espectro (Figura 4.7) el pico de mayor intensidad de 13.325 m/z sólo difiere en 2 Da del peso teórico para la subunidad β de 13.327 Da. En el rango de mayor tamaño la intensidad de los picos disminuye debido a las limitaciones de la técnica, pero se observa que la mayor señal tiene una m/z de 26.430 la cual es 105 Da inferior al peso teórico para la subunidad α de 26.535 Da.

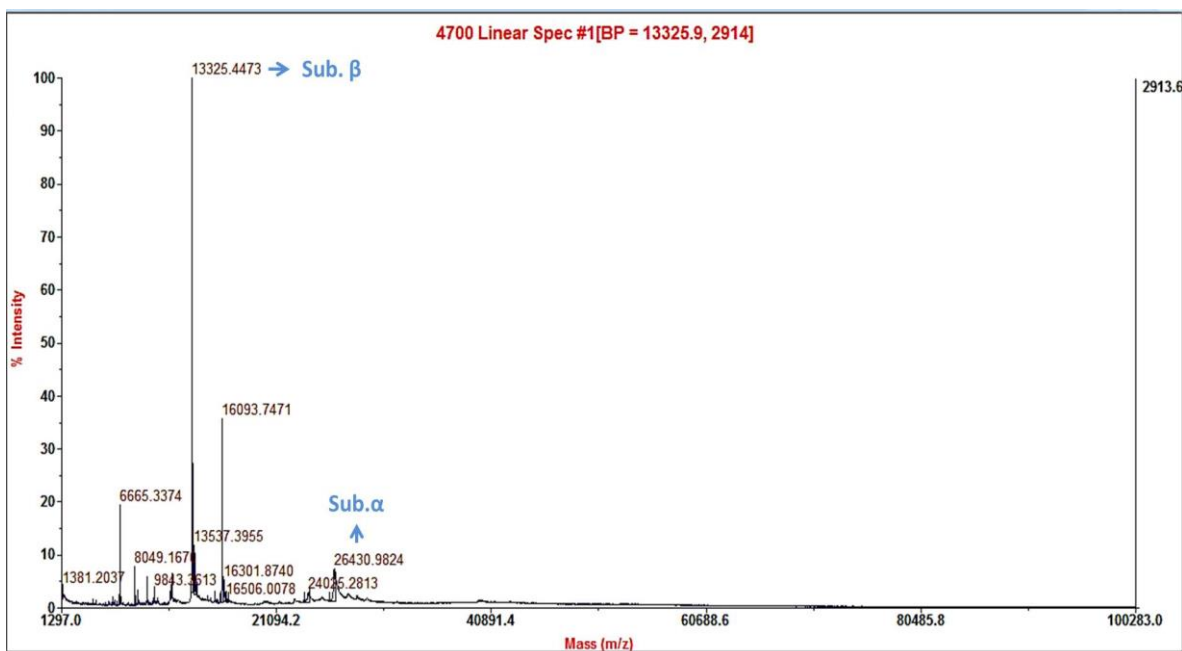


Figura 4.7 Análisis de la estructura primaria de la proteína PpASPG1 mediante espectrometría de masas. Espectro obtenido por MALDI-TOF en modo lineal sobre la proteína purificada. El pico de m/z 13325.4473 corresponde con la subunidad β y el de m/z 26430.9824 con la subunidad α .

Para obtener seguridad sobre la composición de la proteína activa se cargó la muestra en un gel de poliácridamida al 12,5%, los péptidos se separaron por electroforesis SDS-PAGE y se visualizaron por tinción Coomassie, seleccionándose un total de 10 bandas correspondientes a polipéptidos con masa molecular en el rango de 13 a 28 kDa. Todos los péptidos seleccionados se digirieron con tripsina y se analizaron con el equipo MALDI TOF en modo de reflector. De esta manera nos aseguramos detectar todos los componentes de la asparraginasa activa incluso en el caso de que el procesamiento resultara en un acortamiento de la subunidad α . El resultado confirmó que las masas moleculares de 26.430 Da y 13.325 Da corresponden a las subunidades α y β respectivamente. La figura 4.8 muestra la secuencia de la PpASPG1, separada en sus subunidades y se muestran los péptidos identificados mediante MALDI TOF marcados con colores, los cuales cubren un 76,6% de la secuencia de la subunidad α y 60,3% de β . El resultado indica que el sitio de procesamiento

del precursor está conservado en la treonina en posición 245 y la metionina en la posición 1 no está presente en la enzima madura. Esta eliminación reduce el peso teórico de la subunidad α a 26.404 Da y la diferencia de 26 Da con su peso real podría deberse a modificaciones postraduccionales. El resto de polipéptidos analizados pertenecían a contaminaciones presentes en la muestra, principalmente en forma de anhidrasa carbónica.

Para asegurarnos que la secuencia completa de la RV forma parte de la enzima activa, el péptido tríptico derivado del extremo carboxilo de la subunidad α se analizó mediante espectrometría en tándem MALDI TOF/TOF. El espectro de la fragmentación se muestra en la figura 4.9 y confirma que el proceso de activación de la PpASPG1 consiste en un procesamiento que genera dos subunidades sin eliminación de ningún fragmento. Además, el enlace peptídico sustrato de este procesamiento está conservado en la enzima de pino, y por el procesamiento del precursor se genera la Thr245 catalítica en el extremo amino de la subunidad β .

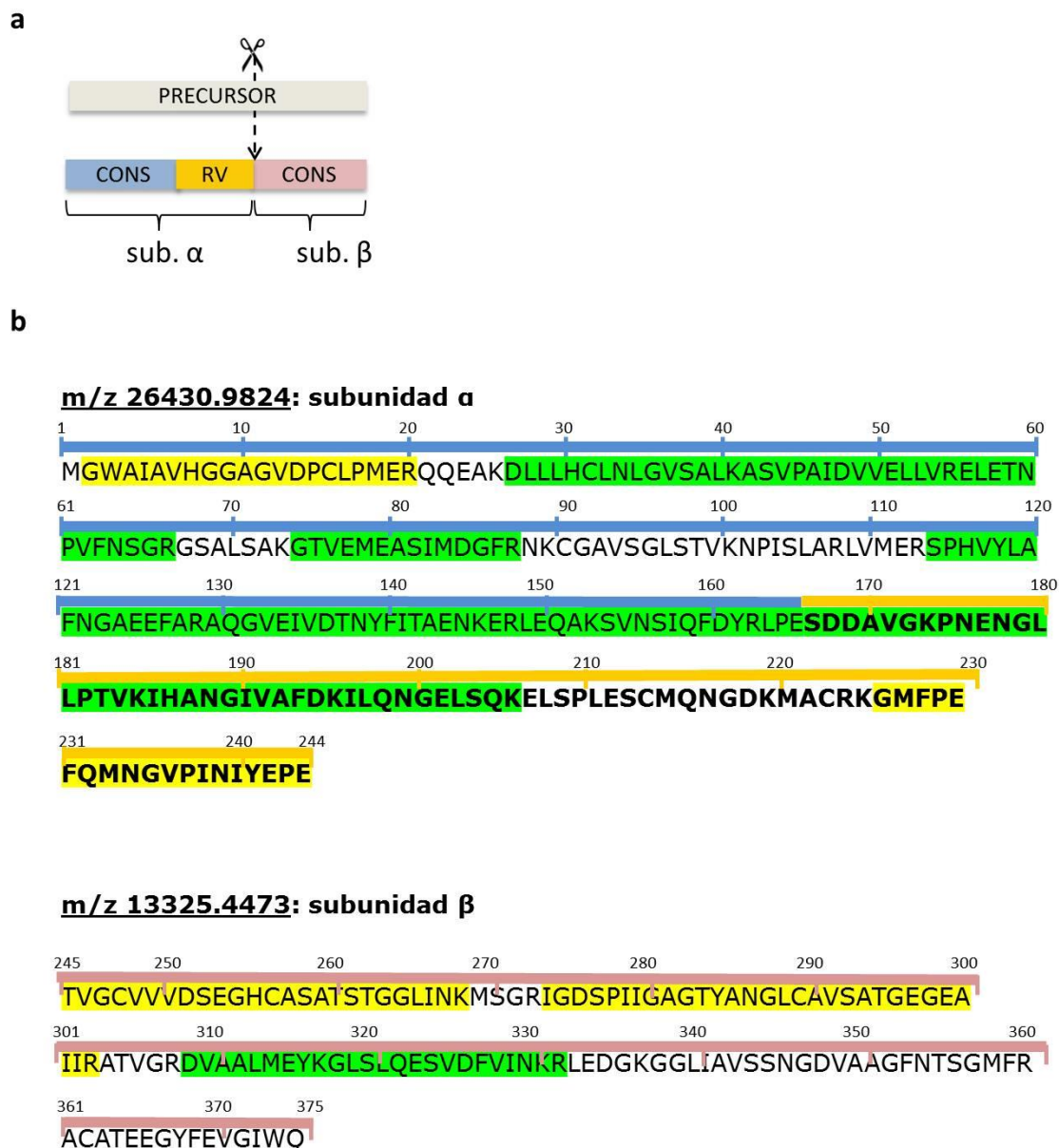


Figura 4.8 Péptidos de PpASPG1 identificados por espectrometría de masas en modo reflector sobre una muestra digerida con tripsina. a) Esquematisación de las subunidades resultantes del procesamiento del precursor, con regiones conservadas entre enzimas ortólogas (CONS) y la región de secuencia variable (RV). b) Los péptidos identificados por MALDI-TOF están marcados sobre la secuencia primaria de la PpASPG1, separada en sus dos subunidades, sombreados en verde los péptidos identificados por MALDI-TOF y en amarillo los péptidos que han sido fragmentados y cuya secuencia se confirmó por espectrometría de masas en tándem MALDI TOF/TOF. La RV está marcada en negrita y la barra que representa la secuencia está coloreada como en el esquema del panel a.

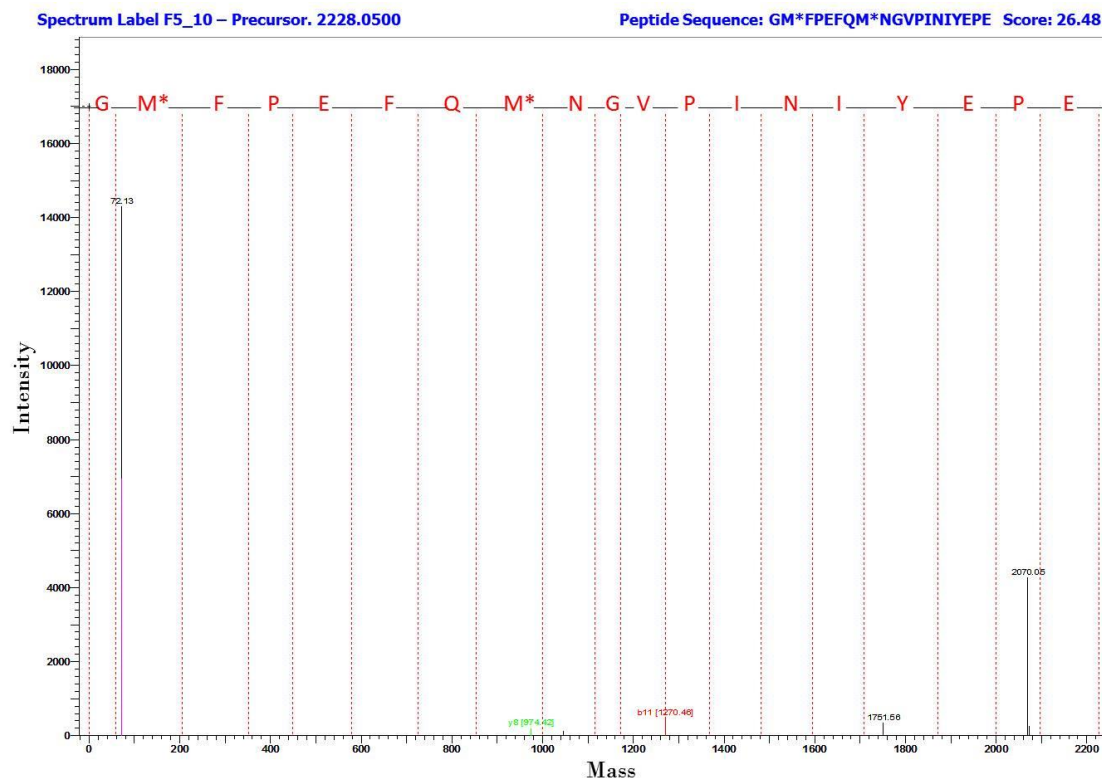


Figura 4.9 Espectro de fragmentación MS/MS del péptido que corresponde al extremo terminal carboxilo de la subunidad α . La secuencia obtenida por MALDI TOF/TOF se lee en letra roja en la parte superior del espectro y coincide con la secuencia del último péptido de la RV. Las metioninas (M*) están sulfoxidadas.

4.3.2 Efecto de alteraciones en la RV sobre el proceso de maduración

Después de haber comprobado que la RV completa forma parte de la proteína madura, nos planteamos averiguar si esta secuencia tiene algún efecto sobre el proceso de la maduración. Para ello hicimos tres variantes de la PpASPG1 de las cuales eliminamos fragmentos progresivamente más grandes de la secuencia que precede al residuo catalítico Thr245. La figura 4.10a representa las construcciones elaboradas para expresar las proteínas recombinantes *in*

planta. Se siguió el mismo procedimiento de agroinfección y el proceso de maduración de los precursores fue analizado en los extractos de las hojas de *N. benthamiana* infiltradas. En la figura 4.10b se muestra la actividad asparraginasa medida en los mismos extractos. Los resultados muestran que la eliminación de distintas porciones de la RV e incluso toda la región RV no suprimió la capacidad del precursor para madurar, ya que detectamos actividad asparraginasa en todos los extractos, aunque el valor obtenido fue considerablemente menor en el caso de PpASPG1-D1. El análisis mediante western blot, mostrado en la figura 4.10c, confirmó la presencia en todos los casos de polipéptidos que coinciden con el peso molecular de la subunidad β , resultado de un correcto procesamiento en todas las variantes. Sin embargo, a diferencia de lo observado para la proteína PpASPG1 no modificada, en los extractos de las versiones mutantes de la proteína se detectó la presencia de ciertas cantidades de precursor sin procesar. Además, se detectaron polipéptidos con un tamaño inferior al del precursor pero superior al tamaño de la subunidad β . Puesto que el antígeno reconocido por el anticuerpo anti-HA se situó en el extremo carboxilo del precursor, se dedujo que estos polipéptidos se corresponden con productos que parecen ser el resultado de procesamientos anómalos del precursor en un enlace peptídico erróneo más hacia el extremo amino del precursor.

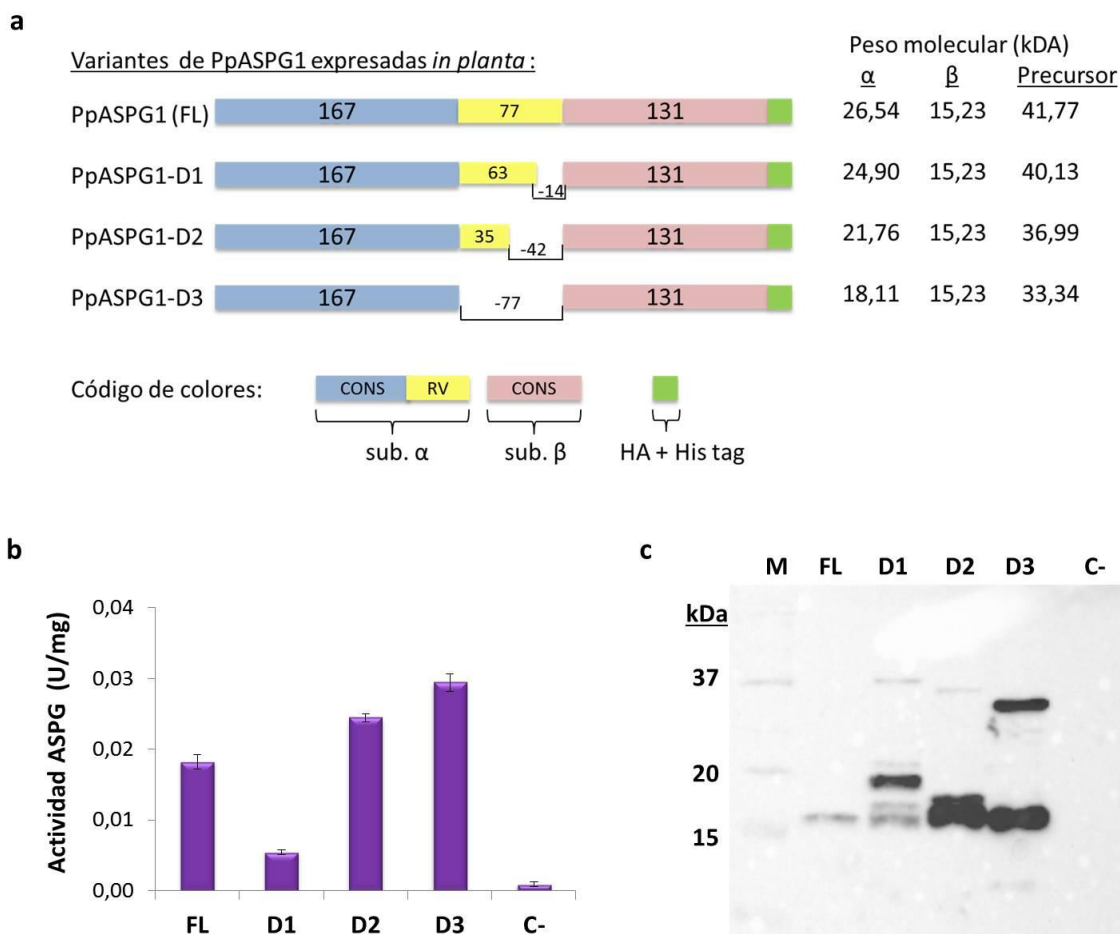


Figura 4.10 Expresión *in planta* de versiones de PpASPG1 mutadas con eliminaciones graduales de la RV. Se utilizó el método de expresión transitoria en hojas de *Nicotiana benthamiana* mediante agroinfección para estudiar el papel de la RV sobre el proceso de la activación de PpASPG1. a) Representación de las diferentes variantes de proteína recombinante expresadas *in planta*. Se eliminaron porciones crecientes de la RV (el número de residuos se marca con signo negativo) y las secuencias codificantes resultantes para la subunidad α se fusionaron a la secuencia codificante de la subunidad β para reconstruir la secuencia codificante del precursor modificada. Todas las construcciones llevan un tag al extremo carboxilo y los pesos moleculares teóricos correspondiendo al precursor y subunidades que resultan del procesamiento se especifican a la derecha. b) Tras la extracción de las proteínas solubles del tejido agroinfiltrado se determinó la actividad con L-Asn 20 mM como sustrato e incluyendo el extracto de una hoja infiltrada con el vector sin inserto clonado como control negativo (C-). Se representa la actividad media de tres replicas en U por mg de proteína total \pm SD. c) Western blot de los extractos de proteínas utilizando un anticuerpo contra el tag localizado en el extremo carboxilo del precursor. En cada calle se analizó 50 μ g de proteínas totales. FL representa la

proteína PpASPG1 completa. D1, D2 y D3, representan las formas modificadas de la proteína por eliminación de 14, 42 y 77 aminoácidos de la RV respectivamente, y que en el texto se denominan PpASPG1-D1, PpASPG1-D2 y PpASPG1-D3.

4.3.3 Capacidad de autoprocesamiento de PpASPG1

Varios autores que han estudiado asparraginasas de especies angiospermas observaron un autoprocesamiento *in vitro* de la proteína recombinante expresada en sistemas heterólogos (Borek *et al.*, 2004; Bruneau *et al.*, 2006; Gabriel *et al.*, 2012). Sin embargo, esto no fue el caso para la ASPG de *Pinus sylvestris* producida en *E. coli*, en la cual sólo se observó procesamiento de la proteína *in vitro* tras la adición de extractos de proteína de pino (Cañas *et al.*, 2007). Según nuestros resultados parece que el proceso de maduración del precursor *in planta* es más eficiente en el caso de la PpASPG1 que en las versiones en las que se eliminó parte o toda la RV. Puesto que las proteínas de angiospermas en las que se observaron el procesamiento *in vitro* presentan una RV de menor longitud a PpASPG1, se decidió evaluar la capacidad de procesamiento *in vitro* de la proteína mutante de PpASPG1 en la que se eliminó completamente la RV (PpASPG1-D3). Procedimos a la purificación de PpASPG1-D3 expresada en *N. benthamiana*, siguiendo el mismo protocolo convencional empleado para purificar PpASPG1. La proteína recombinante se recuperó con un factor de purificación de 10,73 y un rendimiento de 1,81%. Luego, las dos versiones PpASPG1 y PpASPG1-D3 se incubaron durante tres días consecutivos a temperatura ambiente para medir la actividad residual cada 24 horas. La gráfica en la figura 4.11a muestra la actividad específica registrada. La proteína PpASPG1 mantiene una actividad estable durante dos días pero se reduce a un 75% el tercer día, posiblemente como resultado de desnaturalización o degradación. Sin embargo, para PpASPG1-D3, que carece de RV, observamos un aumento continuo de actividad durante todo el tiempo de la incubación lo que asociamos a un posible proceso de maduración *in vitro*. Para comprobarlo, las preparaciones purificadas de ambas proteínas

recombinantes se analizaron mediante western blot (Figura 4.11b). La subunidad α de PpASPG1 fue identificada mediante un anticuerpo que reconoce específicamente a la RV y no se aprecian cambios en la intensidad de la señal, ya que fue completamente procesada *in planta*, como se observó en un análisis previo (Figura 4.10). Por otro lado, para PpASPG1-D3 que se visualiza mediante el anticuerpo contra el péptido HA situado en el extremo carboxilo aparece una banda correspondiente al tamaño de la subunidad β que se va intensificando con el paso del tiempo, lo que confirmó que el aumento de actividad observada fue resultado de la maduración por autoprosesado *in vitro* de parte del precursor PpASPG1-D3.

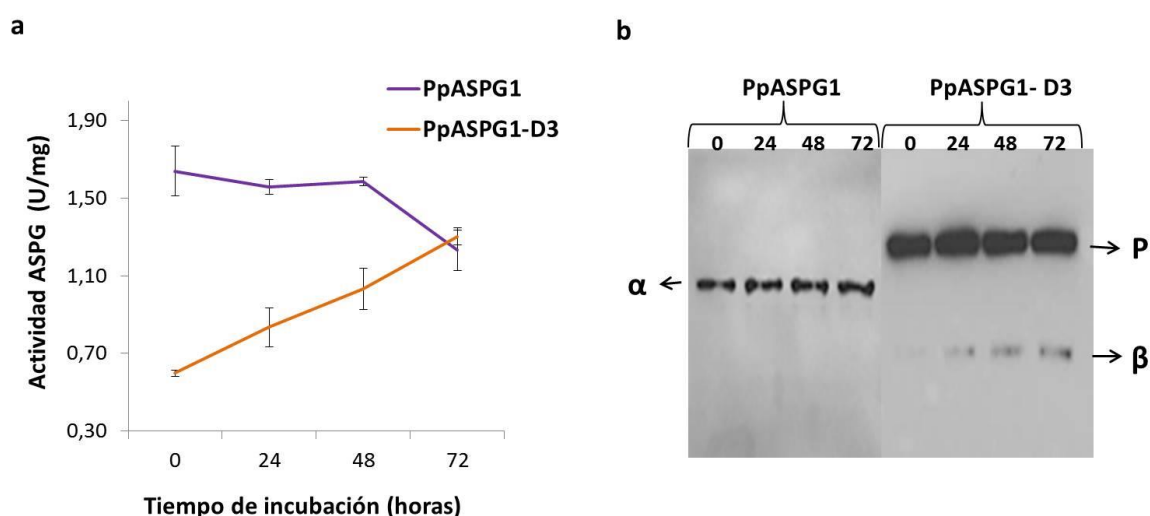
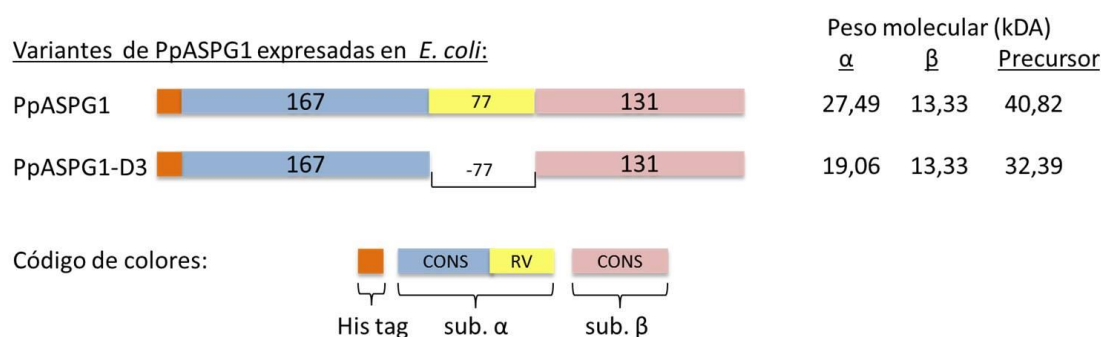


Figura 4.11 Análisis de la capacidad de autoprosesamiento *in vitro* de PpASPG1 y PpASPG1-D3. Las proteínas recombinantes purificadas de hojas de *N. benthamiana* se incubaron durante tres días a temperatura ambiente y el proceso de autoactivación fue seguido por determinación de la actividad residual cada 24 horas y observación del grado de procesamiento mediante western blot. a) Seguimiento durante la incubación de tres días de la actividad específica de ambas proteínas producidas *in planta*. b) Análisis mediante western blot a distintos tiempos de la incubación de las proteínas recombinantes producidas *in planta*. Se empleó un anticuerpo específico para la RV en el caso de PpASPG1 y un anticuerpo con el tag HA para PpASPG1-D3. Cada calle contiene 2 μ g de proteína purificada. P = precursor, α y β = subunidades resultantes del procesamiento.

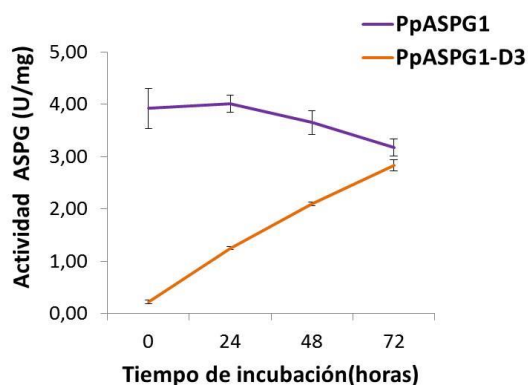
Puesto que la PpASPG1 extraída de hojas de tabaco estaba completamente procesada, decidimos emplear el sistema heterólogo de *E. coli* con el objetivo de obtenerla en forma de precursor y poder comparar la capacidad de autoprosesamiento con la forma PpASPG1-D3. Las dos versiones de proteína recombinante se expresaron en células de *E. coli* mediante el vector pDEST17, que incorpora una extensión de poli-His en el extremo amino del precursor. La figura 4.12a representa un esquema de las dos variantes expresadas en este sistema heterólogo. Se purificaron las proteínas recombinantes en condiciones desnaturizantes a partir de cuerpos de inclusión, usando cromatografía con una resina que incluye iones Ni^{2+} . Las proteínas se recuperaron mayoritariamente en forma de precursor inactivo y la proporción de proteína procesada respecto al total de proteína recuperada fue mayor para PpASPG1 que para PpASPG1-D3, como se observa en la figura 4.12c. Las proteínas recombinantes se incubaron durante tres días como se ha descrito previamente para las expresadas *in planta*, y la figura 4.12b muestra la actividad asparraginasa medida cada 24 horas. A tiempo cero la diferencia en el grado de procesamiento entre las dos formas se traduce en una mayor actividad en PpASPG1, pero mientras pasa el tiempo de incubación la actividad específica medida para PpASPG1-D3 aumenta hasta un nivel similar a PpASPG1 a 72 horas. Aunque la muestra de PpASPG1 también contiene una gran proporción de proteína en forma de precursor, como se observa en el gel de electroforesis SDS-PAGE teñido con plata (Figura 4.12c), esta no se procesa *in vitro* ya que no se produce un incremento en la actividad catalítica ni en la intensidad de las bandas correspondientes a ambas subunidades resultantes del procesamiento. Por el contrario, en el caso de PpASPG1-D3 la tinción de plata muestra una intensificación de las bandas correspondientes a los productos del autoprosesamiento (subunidades α y β), que eran apenas visibles a tiempo cero. La expresión de las proteínas recombinantes en el sistema de *E. coli* confirma que la PpASPG1 se procesa más eficientemente dentro del ambiente celular que la proteína mutada que carece de RV. Sin embargo, la presencia de la RV parece impedir el autoprosesamiento *in vitro*, ya que la proteína no modificada, al igual que la *P. sylvestris*, es incapaz de autoprosesarse *in vitro*,

mientras que su eliminación genera una proteína capaz de autoprocesarse *in vitro*.

a



b



c

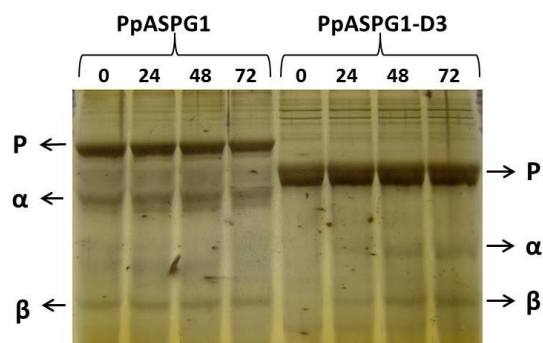


Figura 4.12 Análisis de la capacidad de autoprocesamiento *in vitro* de PpASPG1 y PpASPG1-D3, expresadas en células de *E. coli*. Las proteínas recombinantes purificadas se incubaron durante tres días a temperatura ambiente como descrito para la figura 4.11. a) Esquema representando la estructura de las proteínas recombinantes expresadas en células de *E. coli* y el peso molecular teórico. b) Seguimiento durante la incubación de tres días de la actividad específica de ambas proteínas recombinantes producidas en *E. Coli*. c) Análisis mediante electroforesis SDS-PAGE y tinción de plata de las proteínas expresadas en *E. coli*. P = precursor, α y β = subunidades resultantes del procesamiento. Se cargaron 5 μ g de proteína purificada por pocillo.

4.4 Efecto de la RV sobre las características bioquímica de PpASPG1

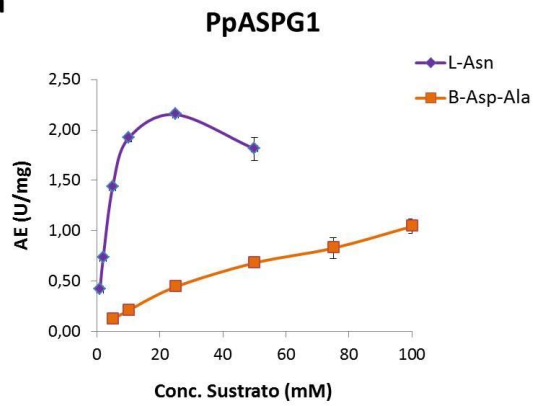
4.4.1 Efecto de la supresión de la RV sobre las características catalíticas de PpASPG1

Las características bioquímicas de un número de asparraginasas de plantas han sido publicadas, pero entre ellas no encontramos características de la enzima de especies gimnospermas. A continuación se describen las características bioquímicas determinadas para PpASPG1 y, en paralelo, se muestran los resultados obtenidos sobre PpASPG1-D3 para comparar posibles diferencias debida a la presencia o ausencia de la región variable en el extremo carboxilo de la subunidad α .

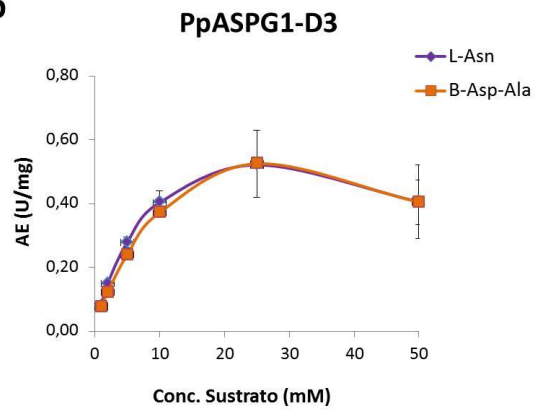
La velocidad de la reacción catalizada por PpASPG1-D3 se mantuvo lineal durante 90 min a 37°C y fue proporcional a la cantidad de enzima añadida. La cinética de la PpASPG1 no modificada ya se describió en la sección 4.2.2, donde es comparada con las otras isoformas de asparraginasa en el pino. En la figura 4.13a se presenta nuevamente los datos para esta enzima presentando la velocidad de la reacción para ambos sustratos en una misma gráfica, con la intención de compararlos con la cinética de PpASPG1-D3 que carece de RV (Figura 4.13b).

Las gráficas presentadas en las figuras 4.13c-f muestran el cálculo de los valores K_m y V_{max} para ambos sustratos, para PpASPG1 y PpASPG1-D3. Como ya se había descrito previamente en la sección 4.2.2, se observan diferencias importantes para ambos sustratos en el caso de la enzima PpASPG1. Sin embargo, la eliminación de la RV en PpASPG1-D3 causa cambios importantes en los valores de la K_m y V_{max} para ambos sustratos. De tal forma, que los valores de estos parámetros cinéticos son más similares para ambos sustratos en el caso PpASPG1-D3.

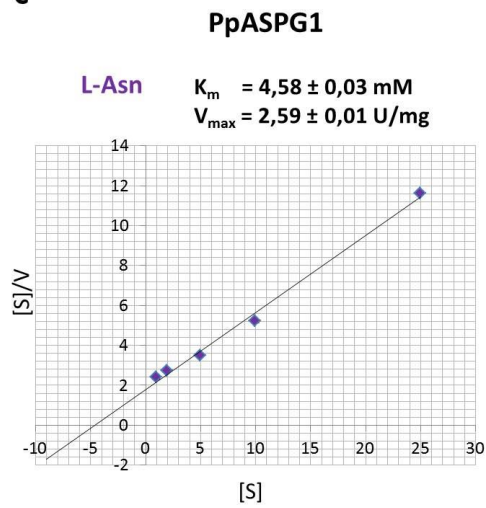
a



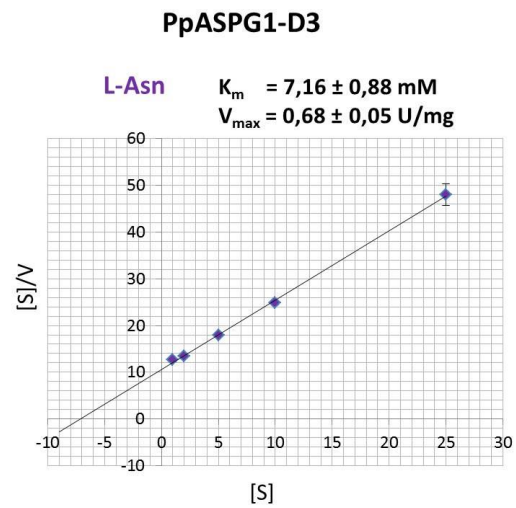
b



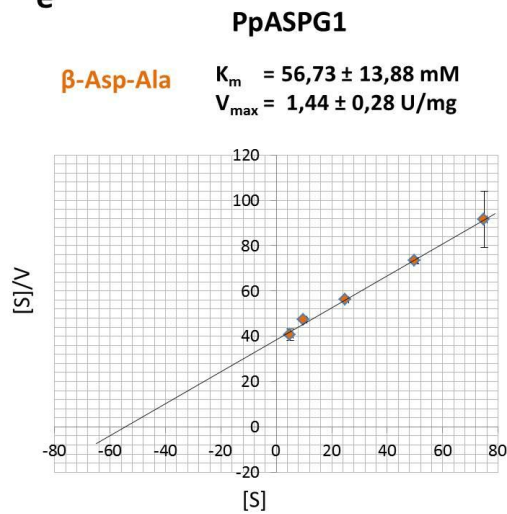
c



d



e



f

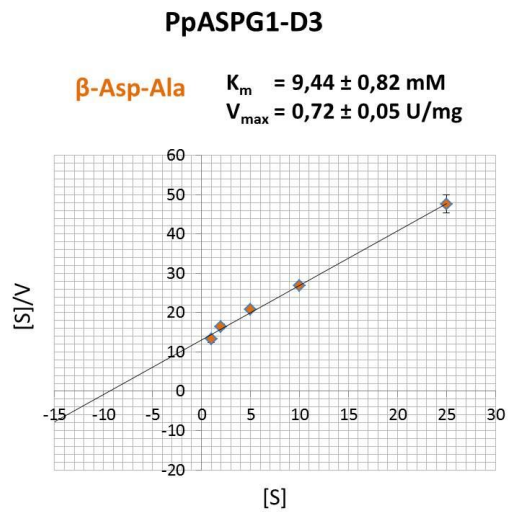


Figura 4.13 Determinación de los valores de K_m y V_{max} de PpASPG1 y PpASPG1-D3 para los sustratos L-Asn y β -Asp-Ala. Se determinó la actividad de ambas proteínas recombinantes a concentraciones crecientes de los dos sustratos (a y b) y se calcularon los valores de K_m y V_{max} mediante una representación de los resultados por el método de Hanes-Woolf para L-Asn (c y d) y β -Asp-Ala (e y f).

En la figura 4.14 se representa la actividad catalítica de PpASPG1 en medios tamponados a diferentes pH. La enzima es activa en un rango entre pH 6 y al menos hasta pH 9 y muestra una velocidad óptima a pH 8. PpASPG1-D3 igualmente es más activa a pH 8, pero parece ser más sensible a pH ácidos ya que a pH 6 no es activa.

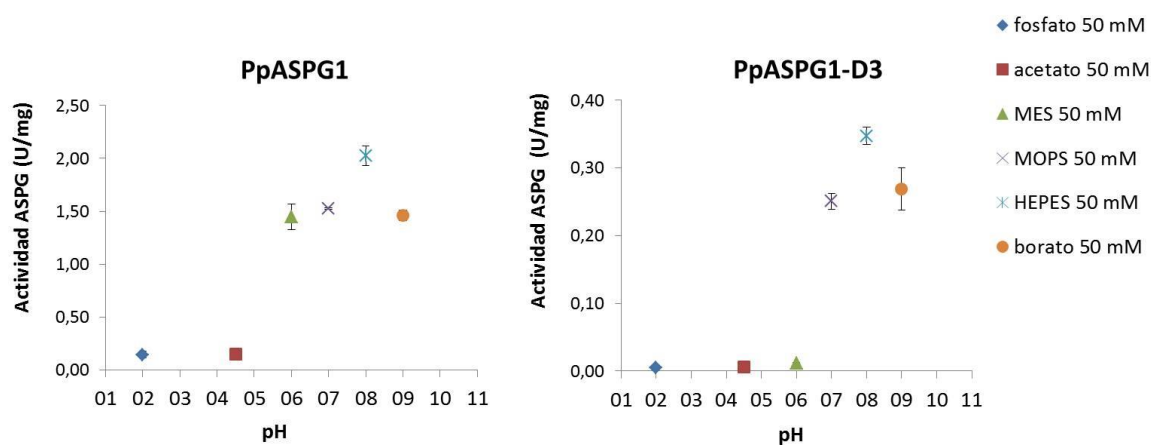


Figura 4.14 Determinación del pH óptimo para PpASPG1 y PpASPG1-D3. Se ensayó la actividad de ambas proteínas con 20 mM de L-Asn como sustrato en diferentes medios tamponados con un pH comprendido entre 2 y 9 para ambas proteínas recombinantes.

Las gráficas 4.15 a y b muestran la velocidad máxima registrada a diferentes temperaturas, en presencia de una concentración saturante del sustrato L-Asn. Se observa una diferencia importante entre la enzima PpASPG1 y PpASPG1-

D3. La reacción de hidrólisis de L-Asn es más eficiente a una temperatura de 35 a 40°C para PpASPG1-D3 y a temperatura mayor disminuye rápidamente. PpASPG1 sin embargo sigue aumentando su velocidad máxima hasta alcanzar un máximo a una temperatura de 50 °C. A partir de estos resultados se calculó la energía de activación para ambas formas de proteína recombinante mediante el ajuste a la ecuación de Arrhenius (Figuras 4.15 c y d). La energía de activación determinada para PpASPG1-D3 de 34,24 kJ/mol es mayor que la de la proteína en su forma nativa de 27,98 kJ/mol, lo que resulta en una menor velocidad de reacción para PpASPG1-D3.

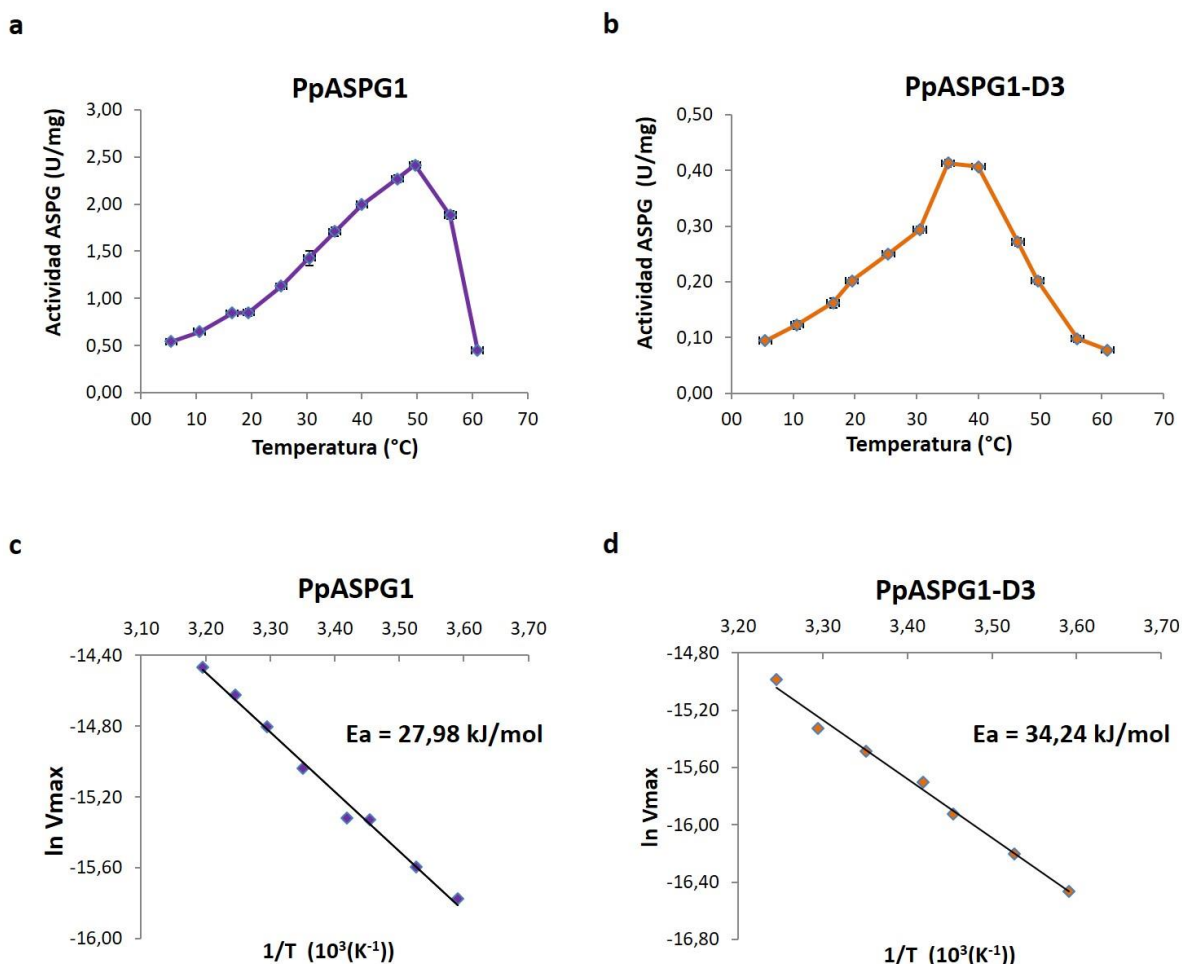


Figura 4.15 Determinación de la temperatura óptima y energía de activación (E_a) para PpASPG1 y PpASPG1-D3. Se ensayó la actividad de ambas proteínas con 20 mM de L-Asn como sustrato a diferentes temperaturas (a y b) y la representación de los valores $1/T$ frente a $\ln V_{max}$ nos permitió calcular la energía de activación para las dos formas de la enzima (c y d).

4.4.2 Contribución de la RV a la estabilidad de la enzima PpASPG1

Después de haber observado que PpASPG1 mantiene su capacidad catalítica a una temperatura tan elevada como 50°C, quisimos obtener más información sobre la estabilidad térmica de esta enzima. Incubamos a 40°C, 50°C, 60°C y 70°C soluciones de las proteínas PpASPG1 y PpASPG1-D3 recombinantes purificadas, retirando alícuotas cada 5 minutos, que se mantuvieron en hielo. La actividad residual de la enzima se determinó y la figura 4.16 muestra los resultados representados como % de la actividad observada a tiempo cero. Se observa que la PpASPG1 es una proteína que resiste una temperatura de 50°C e incluso después de 30 minutos a 60°C mantiene hasta un 75% de su actividad. Por otro lado, la estabilidad térmica de PpASPG1-D3 resulta bastante inferior. La actividad residual tras 30 min a 50°C se ve reducida aproximadamente a la mitad y una temperatura de 60°C provoca la desnaturalización inmediata de la enzima. La temperatura de 70°C necesaria para desnaturalizar PpASPG1, no fue ensayada para PpASPG1-D3.

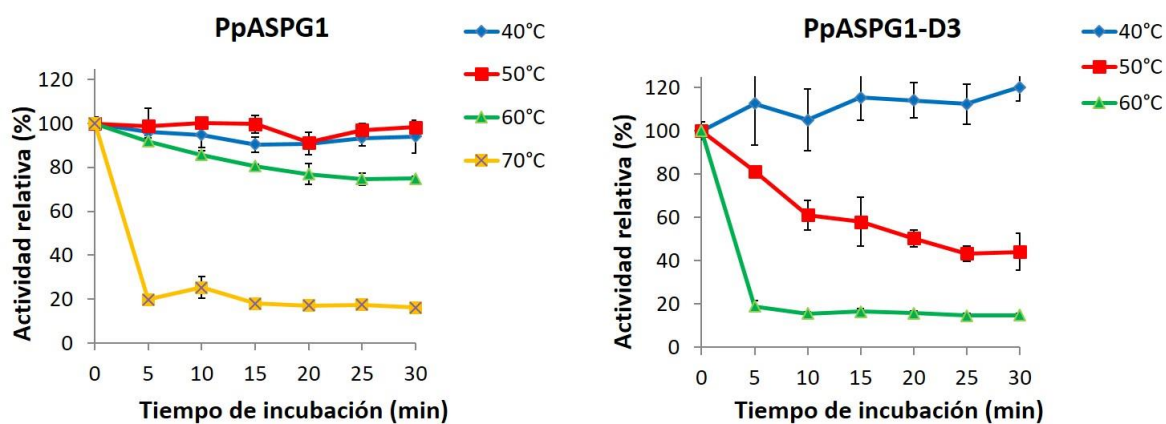


Figura 4.16 Ensayo de la estabilidad térmica de PpASPG1 y PpASPG1-D3 a diferentes temperaturas. Las proteínas recombinantes se incubaron a diferentes temperaturas y se tomaron alícuotas a intervalos de 5 minutos, en las cuales se determinó la actividad asparraginasa. Se representan como % en actividad residual con respecto al control sin incubación previa. Los valores son la media de tres réplicas \pm SD.

4.4.3 Papel del ion potasio en la actividad de PpASPG1 y relación con la RV0,05)

Según el análisis filogenético presentado en el apartado 4.1 la estructura primaria de la proteína PpASPG1 es más similar a la de las asparraginasas dependientes de potasio. Dado que esta clasificación está basada en resultados obtenidos en especies angiospermas se consideró comprobar si el ion tiene la misma importancia para la actividad de esta isoforma en coníferas. Determinamos la actividad asparraginasa en presencia de diferentes iones incluidos en el medio tamponado en forma de sus correspondientes sales con el ion cloruro a una concentración de 50 mM. La figura 4.17a muestra la actividad de la enzima expresada en % con respecto a la mayor actividad registrada a la cual se asignó un valor de 100%. Tanto la enzima PpASPG1 como PpASPG1-D3 presentaban la mayor actividad en el medio con potasio. En presencia de otros iones monovalentes o ninguno (H_2O) la actividad bajó en promedio a un 63% para PpASPG1 y al 75% en el caso de PpASPG1-D3. Para ambas variantes, la diferencia entre la actividad determinada en el medio con potasio y la en los otros medios tamponados fue significativa según el test t de Student ($P \leq 0,05$).

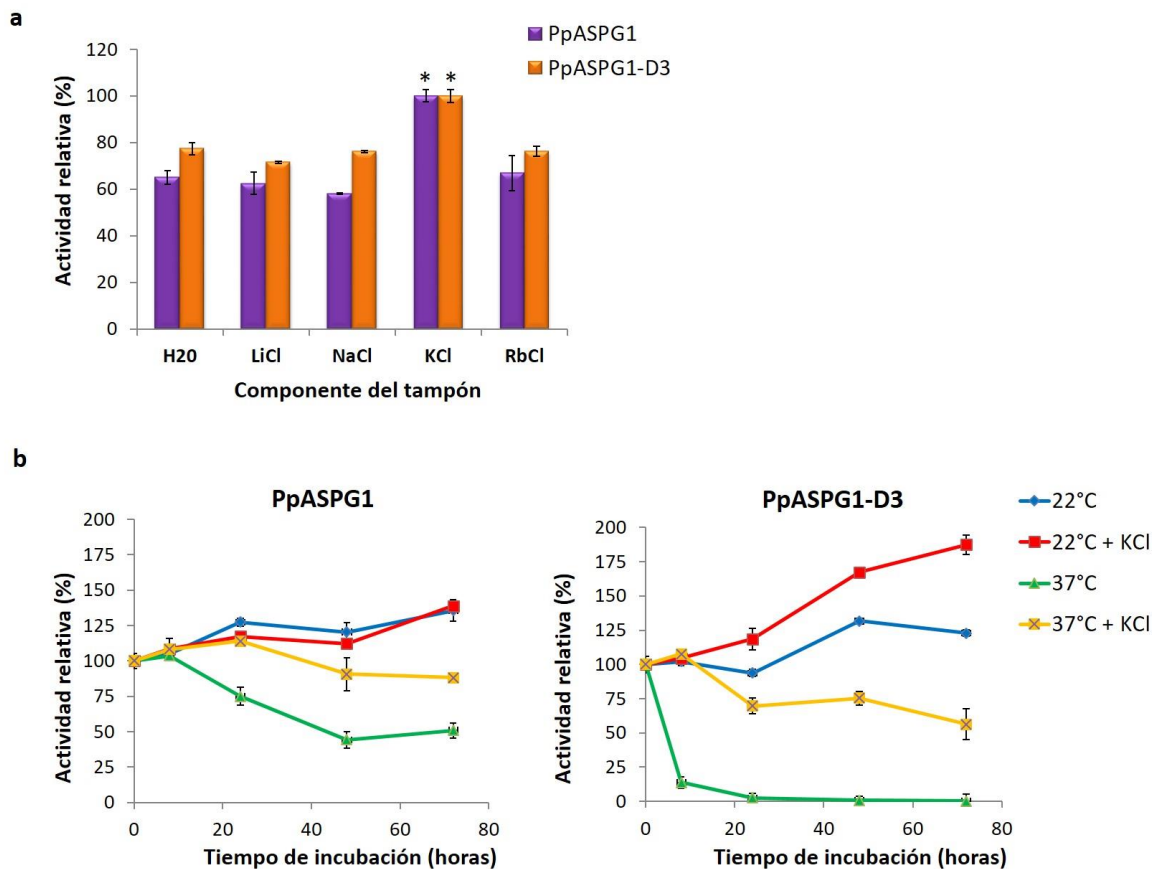


Figura 4.17 Efecto del ion potasio sobre la actividad catalítica y la estabilidad térmica de PpASPG1 y PpASPG1-D3. a) La actividad catalítica de las proteínas recombinantes se determinó en un medio tamponado en presencia de diferentes cationes metálicos monovalentes, en la forma de su correspondiente sal a 50 mM y se incluyó también una reacción sin ningún catión metálico monovalente. La mayor actividad se registró en el tampón con el ion potasio y los otros resultados se presentan como actividad porcentual de ésta. Se realizó una estadística *t* de Student ($P \leq 0,05$) y los valores marcados con * son significativamente diferentes. b) Para estimar el efecto protector del ion potasio contra la inactivación térmica se incubaron alícuotas de las enzimas a 22°C y 37°C, en presencia y ausencia de KCl 50 mM. Se retiraron alícuotas a diferentes tiempos a lo largo de 3 días y la actividad residual fue determinada. Las velocidades medidas son expresadas en % con respecto a muestras no previamente incubadas.

Con el objetivo de averiguar si el potasio tiene un efecto protector frente a la desnaturalización térmica se mantuvieron soluciones de las proteínas recombinantes purificadas a 22°C o 37°C durante tres días, en presencia o ausencia de 50 mM KCl. Se tomaron alícuotas a las 8, 24, 48 y 72 horas y se determinó la actividad residual. En la figura 4.17 b y c se presentan los resultados para las dos proteínas recombinantes (PpASPG1 y PpASPG1-D3), expresados en % de actividad residual en comparación con muestras control que no se sometieron al tratamiento de temperatura previo a la medida de la actividad. La enzima no modificada se mantuvo estable durante al menos tres días a 22°C independientemente de la presencia de potasio, pero a 37°C se observó una mayor disminución de actividad en ausencia del ion. El efecto protector del ion potasio fue más aparente en el caso de PpASPG1-D3. A una temperatura de 22°C la velocidad de reacción de esta enzima aumentó con el paso del tiempo, tanto en presencia como en ausencia del ion potasio, debido al procesamiento *in vitro* de los precursores presentes en la muestra, como se describió en el apartado 4.3.3. Sin embargo, en ausencia del ion potasio el aumento de la actividad catalítica en el tiempo resultó ser menor. A la temperatura de 37°C el efecto protector del ion potasio es más contrastado desde el inicio de la incubación. Así, tras 24 horas a 37°C se mantuvo un 69,76% de la actividad inicial en presencia de potasio, mientras que en su ausencia no se detectó actividad alguna.

4.5 Caracterización de la región promotora de *PpASPG1*

4.5.1 Aislamiento de la región promotora del gen *PpASPG1*

Con objeto de iniciar el estudio de la regulación transcripcional del gen *PpASPG1* aislamos la secuencia promotora mediante una estrategia de Genome Walker. Se hicieron tres réplicas individuales que produjeron fragmentos de promotor de dos tipos. La figura 4.18 es una representación esquemática de la composición de los fragmentos aislados. En dos eventos de amplificación (tipo GW1) la secuencia total de 1633 nucleótidos abarca 1409 pares de bases de región promotora, la región 5'UTR de 129 pb y 95 pb del primer exón del ORF para la asparraginasa (fragmento denominado tipo GW1). Cuando hicimos un tercer evento de amplificación sobre la misma genoteca de DNA con los mismos cebadores, se amplificó un fragmento de tamaño inferior, de 1347 pb, que comparte con el tipo GW1 una región proximal de 437 pb, además de la secuencia 5'UTR y el primer exón del gen. Sin embargo, la región distal del promotor no presenta ninguna similitud con la secuencia del tipo GW1 y además contiene una secuencia de 231 pb que contiene un fragmento parcial del ORF codificante para la malato deshidrogenasa (MDH) de *Pinus pinaster* (fragmento que denominamos tipo GW2).

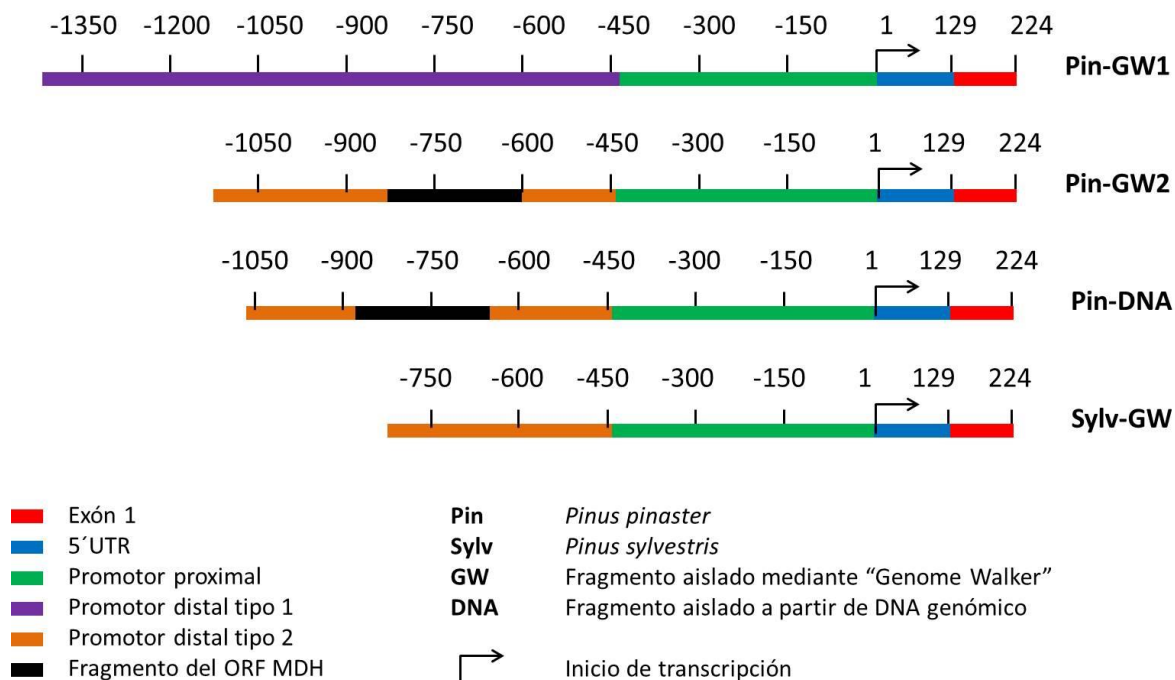


Figura 4.18 Representación esquemática de las secuencias promotoras de *PpASPG1* aisladas.

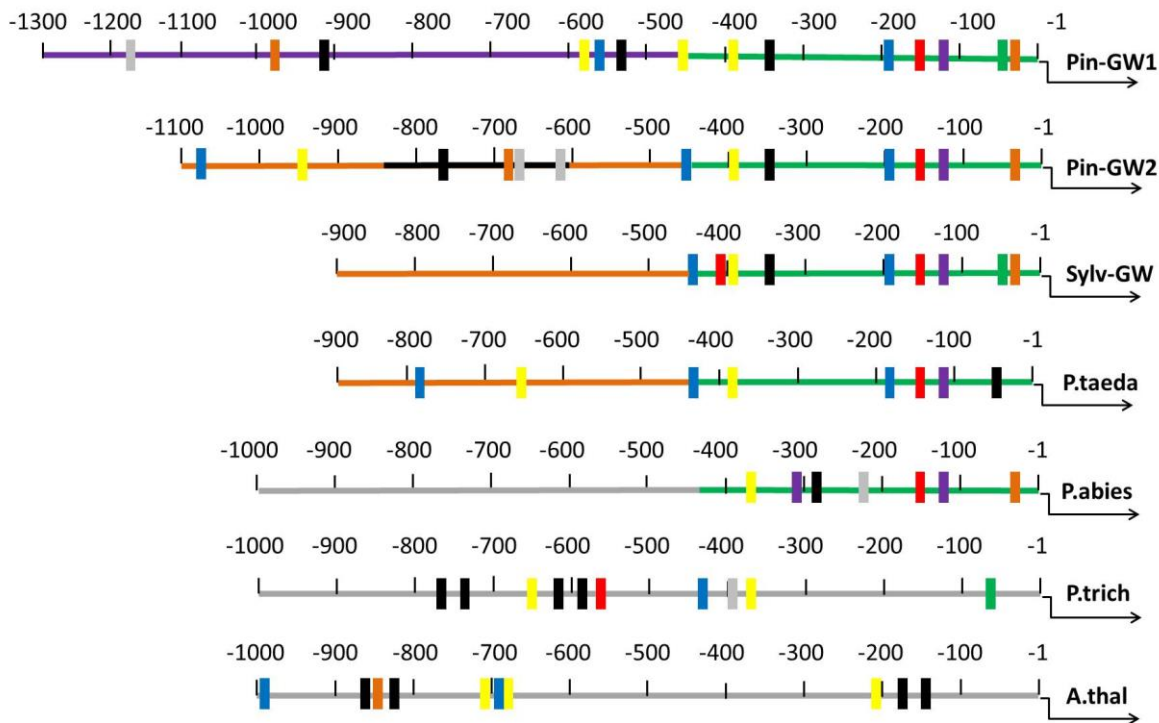
Mediante la técnica Genome Walking se aislaron en tres eventos independientes un total de dos fragmentos de promotor diferentes en *P. pinaster*, denominados Pin-GW1 y Pin-GW2. Las secuencias comparten entre ellas la región proximal, el 5'UTR y una fracción del primer exón del ORF delimitada por el cebador reverso. Estas secuencias están también presentes en un fragmento de DNA genómico aislado directamente a partir de DNA genómico de *P. pinaster* (Pin-DNA) empleando cebadores derivados de la secuencia de Pin-GW2, y otro aislado mediante Genome Walking en *P. sylvestris* (Sylv-GW). Para facilitar su identificación en el esquema las secuencias compartidas se marcan con los mismos colores. La región distal del promotor difiere entre los amplificados Pin-GW1 y Pin-GW2, teniendo la secuencia distal de Pin-GW2 más similitud con las secuencias distales de Pin-DNA y la del promotor de *P. sylvestris* Sylv-GW, aunque este último no presenta la inserción de un fragmento del ORF de Malato deshidrogenasa (MDH). Las distintas posiciones de los nucleótidos de la región promotora están marcadas con signos negativos, mientras que los de la secuencia transcrita se marcan con números positivos.

Consideramos la posibilidad de que el fragmento aislado tipo GW2 fuera un artefacto de la técnica aplicada, producto de la ligación de fragmentos no contiguos en el genoma durante la síntesis de las genotecas. Para comprobar la existencia de estas secuencias en el genoma del pino se diseñaron cebadores en dirección sentido sobre las secuencias promotoras aisladas de los dos tipos y se extrajo DNA genómico de los cotiledones de plántulas germinadas a partir de semillas originando de la Sierra de Alcaraz (España). La combinación del cebador tipo GW2 con el cebador de la secuencia localizada en el exón de la asparraginas produjo un fragmento de 1283 pb (Pin-DNA) casi idéntico a la secuencia aislada tipo GW2, igualmente presentando la fracción codificante para MDH aunque desplazada 57 pb más hacia 5'. Una secuencia promotora de ASPG de 823 pb previamente aislada por nuestro grupo de investigación de *Pinus sylvestris* (Sylv-GW) comparte también la región proximal del promotor y presenta mayor similitud con el fragmento GW2 que GW1, con la excepción de que el promotor de esta especie no incluye el fragmento que codifica MDH. Con los cebadores específicos del extremo 5' del fragmento tipo GW1 no obtuvimos una buena amplificación sobre el DNA genómico después de varios intentos.

4.5.2 Análisis *in silico* de la secuencia promotora

Las secuencias de promotores aisladas se analizaron *in silico*, mediante rastreos en bases de datos de elementos *cis* de promotores de plantas. El esquema de la figura 4.19 representa la posición dentro de la región promotora de los posibles elementos reguladores *cis* más relevantes y para su comparación se muestran también los resultados sobre las regiones promotoras de asparraginasas de otras coníferas (*P. sylvestris*, *P. taeda* y *P. abies*) y dos especies modelo de angiospermas (*A. thaliana* y *P. trichocarpa*). El esquema ilustra el alto grado de conservación de la región promotora proximal de ASPG en coníferas. La región distal por otro lado apenas presenta homología en cuanto a la ubicación de los sitios *cis* y aunque el promotor de *P. sylvestris* tiene una mayor similitud al promotor tipo GW2 en la región distal, las diferencias en nucleótidos existentes resultan en diferencias en elementos reguladores *cis* según las bases de datos consultadas. La distribución de los elementos reguladores en especies angiospermas no tiene similitud con las coníferas, pero varias de las mismas cajas están presentes en regiones más distales del promotor. El análisis *in silico* muestra en el promotor proximal la presencia de elementos *cis* relacionados con respuestas a las hormonas ABA y jasmonato, además de a algunas condiciones de estrés, que causan cambios en los niveles de transcritos del gen en plántulas como se describe en el apartado 4.6.3.

En la región distal del promotor, la inserción del fragmento ORF para MDH en la secuencia GW2 curiosamente introduce tres elementos *cis* que también presenta el promotor distal GW1. Aparte de las cajas presentadas en la figura el análisis *in silico* identificó numerosos posibles elementos relacionados con un control por luz, factores de transcripción, más factores de estrés y fitohormonas las cuales por razones de claridad no se han incluido en el esquema. Con la presencia de una caja TATA en posición -30, la secuencia aislada reúne los requisitos mínimos para constituir un promotor funcional.



Elementos *cis*

- | | |
|---|----------------------------------|
| ■ Inducción por citoquinina | ■ Respuesta a baja temperatura |
| ■ Respuesta a jasmonato | ■ Expresión circadiana |
| ■ Inducción por ácido salicílico (WRKY) | ■ Inducción por sal y patógenos |
| ■ Respuesta a ABA | ■ Respuesta a patógenos y estrés |

P.taeda	<i>Pinus taeda</i> (PITA_000017050)
P.abies	<i>Picea abies</i> (MA_10435256g0010)
P.trich	<i>Populus trichocarpa</i> (POPTR_0014s02300)
A.thal	<i>Arabidopsis thaliana</i> (At3g16150)

Figura 4.19 Posición de elementos *cis* en las regiones promotoras de asparraginasas de varias especies. Los dos tipos de secuencias de promotor aislados de *P. pinaster* (Pin-GW1 y Pin-GW2) y de *P. sylvestris* (Sylv-GW) se analizaron *in silico* para buscar elementos regulatorios *cis* y se compararon con los ortólogos de algunas otras especies. El esquema representa las posibles cajas reguladoras más relevantes con un código de colores y sus localizaciones en cada fragmento amplificado respecto a la secuencia de nucleótidos indicada con números en la parte superior de cada promotor. La barra que representa el fragmento de promotor mantiene los mismos colores como la figura 4.18 para las regiones que presentan más similitud entre ellas.

4.5.3 Análisis funcional de la región promotora

Para comprobar la actividad promotora de las secuencias aisladas y arrojar luz sobre la funcionalidad de ambas variantes se realizó un ensayo funcional en un sistema heterólogo, consistente en hojas de *N. benthamiana* transformadas con una construcción para testar si la expresión del gen delator luciferasa es dirigida por los promotores tentativos del gen *PpASPG1*. El esquema de la figura 4.20a resume las secuencias que se introdujeron en el vector pBI121, en posición 5' respecto a la región codificante de luciferasa. En el panel b de la misma figura se muestra la actividad luciferasa detectada por mg de proteínas solubles en los extractos de las hojas infiltradas. En tres réplicas independientes la secuencia completa de la región promotora tipo GW1 produjo el triple de expresión del gen delator que la secuencia tipo GW2 y hasta siete veces más actividad que sólo el promotor proximal. En el control positivo se detectó una luminescencia más de un orden de magnitud mayor al observado con la construcción que incluía el fragmento GW1. Extractos de hojas no infiltradas no produjeron luminescencia (datos no presentados). También analizamos la luminescencia relativa por mg de proteína soluble en extractos de *Agrobacterium tumefaciens* previo a la infiltración de las plantas. La presencia de un promotor Lac en el vector pBI121 en posición 5' a la localización del promotor induce una expresión del gen delator en las células de *A. tumefaciens* (Chen and Winans, 1991). La actividad luciferasa en *Agrobacterium*, dirigida por este promotor Lac, es mayor para la construcción con el promotor proximal inserto que para las otras construcciones, como se observa también en la figura 4.20b. La mayor longitud en nucleótidos de promotores insertos en las otras construcciones produce una menor expresión luciferasa en *A. tumefaciens*.

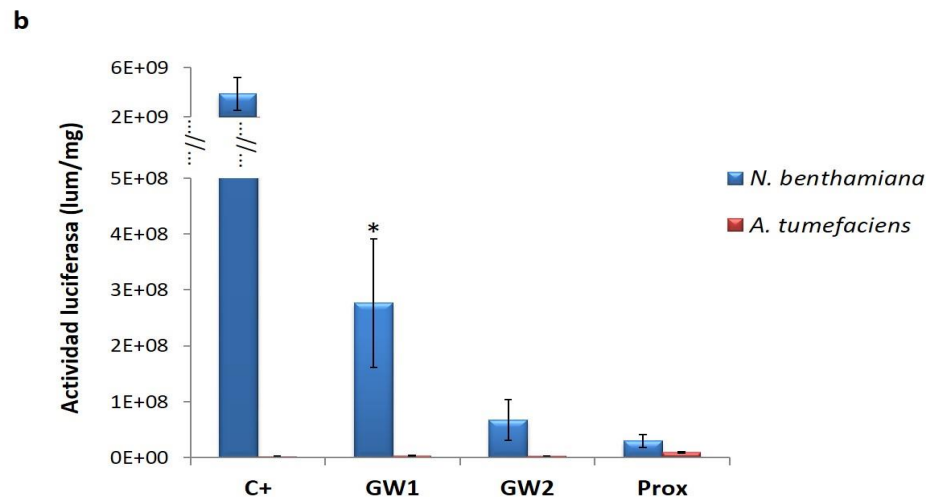
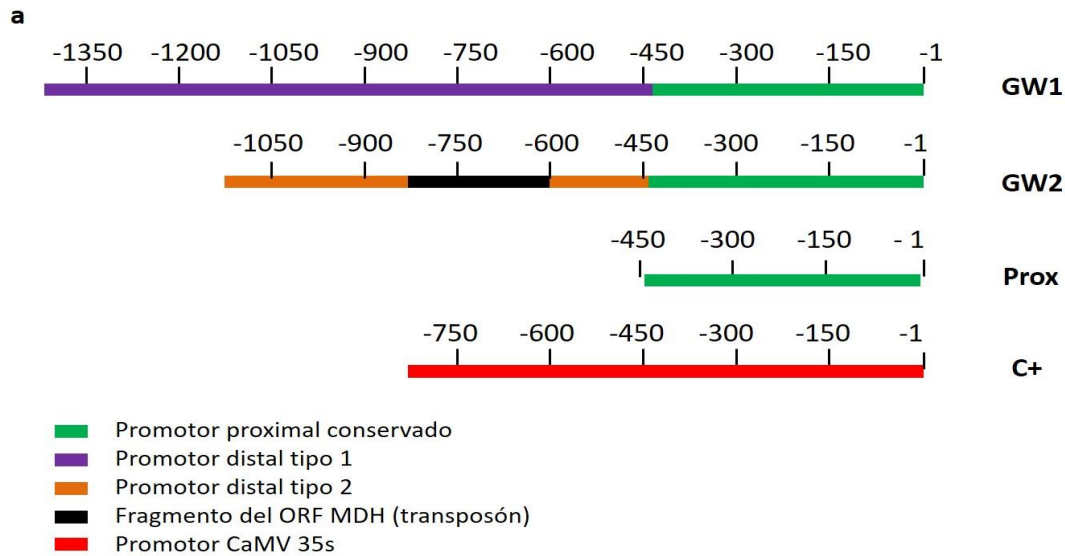


Figura 4.20 Análisis funcional de las regiones promotores de *PpASPG1* por agroinfiltración en hojas de *N. benthamiana*. a) Representación esquemática de los fragmentos de promotor clonados en el vector pBI121 para dirigir la expresión del gen delator de luciferasa. Se hicieron construcciones con las secuencias completas de los dos tipos amplificadas por Genome Walker (GW1 y GW2) y una versión más corta que sólo contiene la secuencia conservada proximal. Como control positivo se introdujo el promotor 35s del virus de mosaico de coliflor (CaMV). b) Tres días después de la agroinfiltración se determinó la luminescencia por mg de proteínas totales en los extractos de tejidos infiltrados de *N. benthamiana*. Se realizó una estadística *t* de Student ($P \leq 0,05$) y los valores marcados con * son significativamente diferentes. La actividad luciferasa detectada en los cultivos de *A. tumefaciens* se presenta como control negativo del experimento y para todas las construcciones fue significativamente inferior a la actividad determinada en las plantas.

Para evitar alteraciones en la eficiencia de expresión *in planta*, hemos preferido no manipular la secuencia del promotor Lac en el vector y utilizar la actividad debido a la expresión en *Agrobacterium* como control negativo. En tres réplicas independientes, la actividad luciferasa en *N. benthamiana* dirigida por los promotores GW1 y GW2 superó a la observada para la construcción que porta el promotor proximal, aunque según el test estadístico *t* de Student la diferencia entre GW2 y el promotor proximal no es significativa ($P \leq 0,05$). En *Agrobacterium* es en el promotor proximal que se detecta mayor actividad del gen delator y para todas las construcciones la actividad luciferasa es significativamente inferior a la registrada en las plantas. Estos resultados indican que el promotor aislado tipo GW1 debe contener en su región distal un(os) elemento(s) que aumentan el nivel de expresión del gen delator *in planta*.

4.5.4 Variabilidad genómica en los promotores de *PpASPG1* de diferentes poblaciones de *P. pinaster*

Los pinos tienen una gran capacidad de adaptación (Sanchez-Gomez *et al.*, 2010) y como la asparraginasa parece tener un papel en la adaptación del árbol a las condiciones internas y externas (ver sección 4.6), cambios en la regulación de este gen podría determinar modificaciones en la capacidad de respuesta de distintas poblaciones. Para obtener una idea sobre la variabilidad genética del gen de asparraginasa aislamos las secuencias promotoras de *PpASPG1* en pinos de 11 procedencias diferentes. Sobre DNA extraído de sus acículas amplificamos las regiones promotoras usando cebadores específicos sobre las secuencias aisladas previamente en combinación con los cebadores antisentido sobre el primer exón del gen. Varios intentos empleando como cebador en 5' oligonucleótidos derivados de las secuencias de la región distal no produjeron ninguna amplificación y sólo obtuvimos fragmentos más cortos con cebadores diseñados sobre la zona del promotor proximal, común entre los tipos GW1 y GW2. Suponemos que la región distal de este gen presenta mucha variación entre poblaciones o incluso entre individuos. Requiere de más

experimentos y/o diferentes métodos para averiguar si esto es el caso y hasta que nivel la región distal influye en la expresión del gen. La figura 4.21 presenta los SNPs encontrados entre las secuencias de promotor proximal de 437 pares de bases, la región 5'UTR y el primer exón del ORF para PpASPG1. Dentro del promotor hay un total de 5 nucleótidos que presentan variación entre las poblaciones.

La variación de G a C mostrada en el mapa número 1 provoca la pérdida de un elemento *cis* de respuesta rápida a deshidratación (*erf*) en la población de la Sierra de Alcaraz, cambio que comparte con la secuencia aislada tipo GW2 y un alelo de un individuo de Puerta de la Vega. El mapa número 2 indica en color naranja las poblaciones que por el cambio de C a A han perdido un elemento regulatorio de función desconocida, un cambio también encontrado en un alelo en las poblaciones marcadas en color negro. El cambio presentado en el mapa nr 3 de G a C en los árboles de origen francés introduce una caja de respuesta a jasmonato en la cadena antisentido con una secuencia central CGTCA. El SNP representado en el mapa número 4 donde G pasa a A en la población de Sierra de Alcaraz (y el evento Genome Walker tipo 2) hace que se pierda la caja *cis* de respuesta a la hormona ABA, como se observa también en la figura 4.20 en la secuencia GW2.

El primer nucleótido del transcrito tiene variación entre las poblaciones y en esta posición también se localiza un elemento de regulación *cis* con función desconocida sólo presente en las poblaciones marcadas con color morado en el mapa número 5. Dentro del 5'UTR hay otro nucleótido que es diferente en el individuo de la Sierra de Alcaraz (mapa número 6) que comparte este cambio con la secuencia GW2. En la población de Oria marcada de color naranja en el mapa número 7 detectamos un cambio en la región codificante que se traduce en el cambio Gln23 → Lys, también presente en la secuencia aislada GW1. La totalidad de 5 SNPs en los promotores proximales de las poblaciones podría tener un efecto sobre la regulación transcripcional de la ASPG en pino, lo cual

podría estar relacionado con las distintas capacidades adaptativas de las diferentes procedencias.

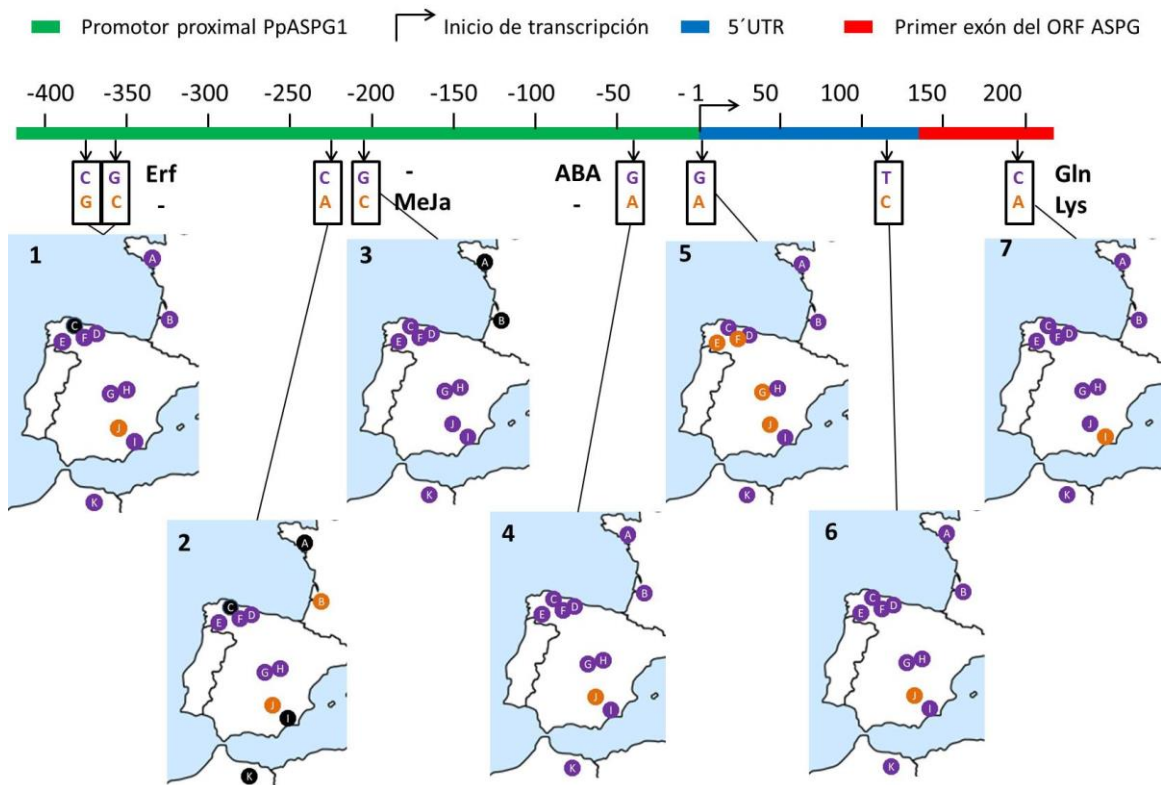


Figura 4.21 Variación genética en la secuencia proximal del promotor de *PpASPG1* en individuos de diferentes procedencias. Se amplificó mediante PCR fragmentos genómicos que incluían la región proximal de los promotores, el 5'UTR y el primer exón del gen, a partir de DNA genómico extraído de plántulas de 11 poblaciones de *Pinus pinaster*, distribuidas entre Francia, España y Marruecos (ver figura 3.2 para más detalle). El esquema representa las secuencias de los fragmentos de DNA aislado y las flechas indican las posiciones de SNPs, marcándose los nucleótidos de las posiciones polimórficas en violeta o naranja. En los mapas conectados con los SNPs las poblaciones se colorean correspondiendo al color del nucleótido que presentan y en las poblaciones marcadas en negro se encontraron ambas variantes en diferentes eventos de secuenciación. Los distintos mapas están etiquetados con números en la parte superior para una fácil referencia a los mismos en el texto. Algunos posibles elementos *cis* regulatorios son afectados por los SNPs (Erf = elemento de respuesta a deshidratación, MeJA = elemento de respuesta a metil-jasmonato, ABA = elemento de respuesta a ABA).

4.6 Perfiles de expresión de los genes de asparraginasa de pino

4.6.1 Perfiles transcripcionales de los tres genes de asparraginasas durante las fases de desarrollo temprano del pino

Con el objetivo de ampliar nuestra información sobre la función de las asparraginasas en el desarrollo del pino se obtuvieron los perfiles de abundancia relativa de mRNA de los tres genes en plántulas de diferentes estadios de desarrollo. Los cotiledones, hipocótilos y radículas se separaron y la cantidad relativa de transcritos de los tres genes se determinó mediante qRT-PCR y se presenta en la figura 4.22. *PpASPG3* es el gen que presentó niveles más bajos de mRNA en todos los órganos y estadios. En cuanto al gen *PpASPG2*, la cantidad de transcritos aumentó gradualmente en el cotiledón, mientras que en el hipocótilo y la radícula sólo se observó un aumento en el estadio de desarrollo más avanzado, duplicándose el nivel de mRNA respecto a los estadios anteriores. Este aumento con el desarrollo también se observa para el gen *PpASPG1* en cotiledones, y especialmente en hipocótilo donde la abundancia en estadio E5 es el triple del estadio E2. Los hipocótilos de ambos estadios se diferencian por pasar de un tejido verde fotosintéticamente activo y con un limitado desarrollo vascular a un tejido con un sistema vascular más desarrollado que empieza a organizar los haces vasculares en la característica estructura cilíndrica (Cañas *et al.*, 2007). En las radículas este gen tiene un alto nivel de transcritos ya desde el primer estadio de desarrollo, hasta 10 veces mayor que en cotiledones. Un análisis estadístico mostró diferencias significativas entre órganos y entre estadios de desarrollo para los tres genes.

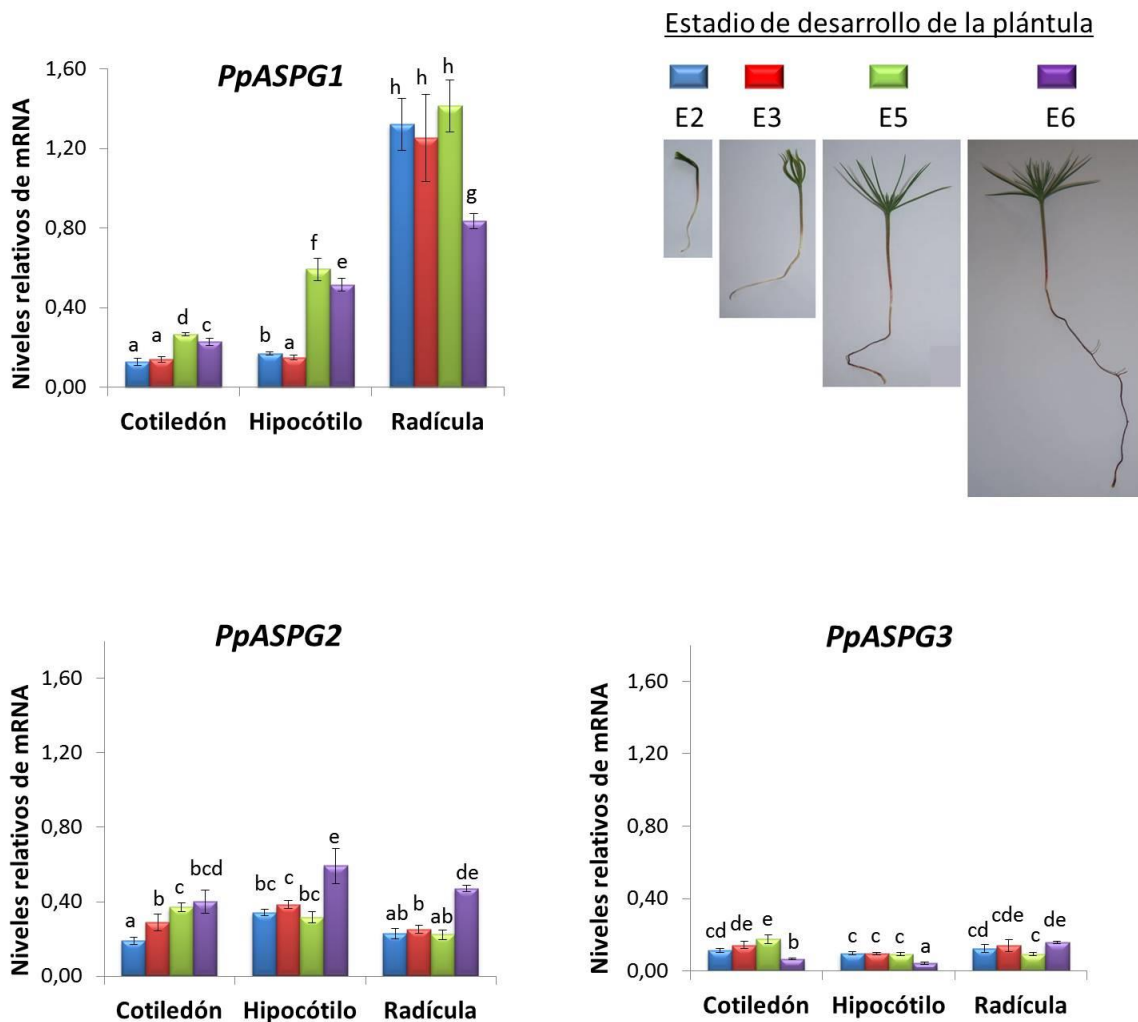


Figura 4.22 Niveles relativos de transcritos de los tres genes de asparraginasa de pino en diferentes estadios de desarrollo de plántulas. Los niveles relativos de mRNA se determinaron mediante qRT-PCR en tres órganos y se representan las medias de tres réplicas \pm SD. Los estadios de desarrollo se determinaron según la morfología descrita en la tabla 3.1. Se realizó una estadística *t* de Student ($P \leq 0,05$) y los valores marcados con letras distintas son significativamente diferentes.

Dado el alto nivel de mRNA del gen *PpASPG1* observado en las radículas se decidió analizar la distribución de los mensajeros para los tres genes en diferentes partes de este órgano, para obtener una imagen más detallada de su distribución. Se emplearon radículas de plántulas en el estadio E4, las

cuales se seccionaron en trozos de 1 cm y también se analizaron las radículas de plántulas en estadio E0. Como se observa en la figura 4.23, el nivel de transcritos del gen *PpASPG1* en la parte más desarrollada (R2, R3, R4) de la radícula es 6 veces mayor al nivel en el ápice de la radícula (R1). Los niveles más bajos se observaron en la radícula recién germinada del estadio E0 (G), un tejido con aspecto similar a la muestra R1.

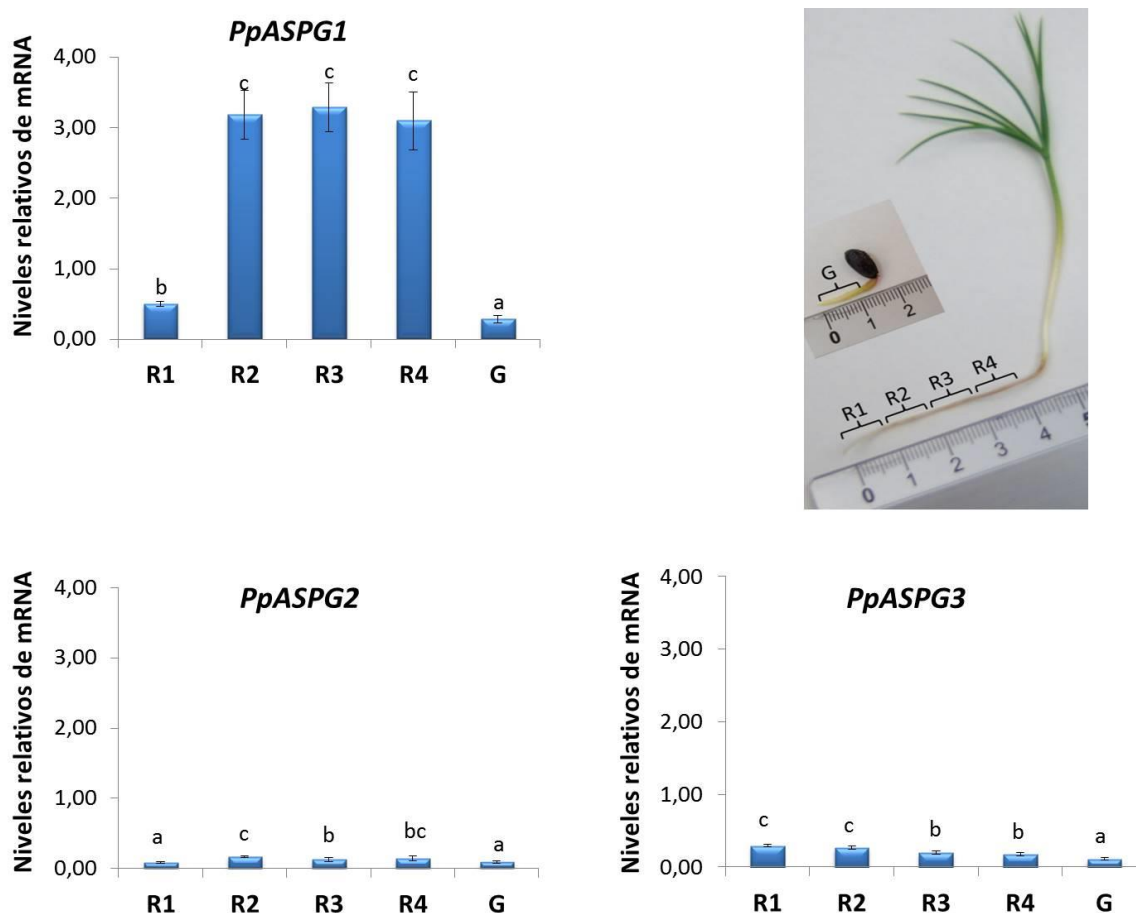


Figura 4.23 Abundancia relativa de transcritos de los tres genes de asparraginasa a lo largo del eje vertical de las radículas de plántulas del estadio E4 y en radículas del estadio E0. Las radículas se seccionaron en trozos de 1 cm a partir del ápice (R1 hasta R4) y de las plántulas en estadio E0 se recolectó 1 cm de la radícula emergida (G). Los niveles relativos de los transcritos en las diferentes partes de tejido se determinaron por qRT-PCR y representan valores medios de 3 réplicas \pm SD. Valores significativamente diferentes están marcados con letras distintas.

Por otro lado, en el caso del gen *PpASPG2*, el patrón de mRNA fue similar al de *PpASPG1*, pero las diferencias en los niveles relativos entre el ápice o la radícula recién germinada y las regiones más desarrolladas de la radícula fueron menores. Finalmente, el nivel relativo de mRNA del gen *PpASPG3* disminuyó ligeramente desde el ápice a la parte superior más desarrollada de la raíz, observándose los niveles más bajos en la radícula recién germinada.

Puesto que el nivel de transcritos de *PpASPG1* y *PpASPG2* aumentó en los estadios más avanzados del desarrollo del hipocótilo, se analizaron los niveles de transcritos de ambos genes en tallos de pinos con una edad de 2 años, para determinar si estos genes siguen desempeñando una función en tallos y si su expresión está asociada al grado de diferenciación del mismo. Para ello, dividimos los tallos de tres individuos en tres partes como se muestra en la figura 4.24. En la foto se diferencian morfológicamente las fracciones de la planta de nuevo crecimiento "brote", del año anterior "tallo superior" y del primer año "tallo inferior". Se observa que ambos genes se expresan en las tres secciones de tallo analizadas. En el caso del gen *PpASPG1* los niveles de transcrito fueron similares en las tres secciones, mientras que en el caso del gen *PpASPG2* se apreció un ligero incremento de los niveles de mensajero desde el brote a la base del tallo.

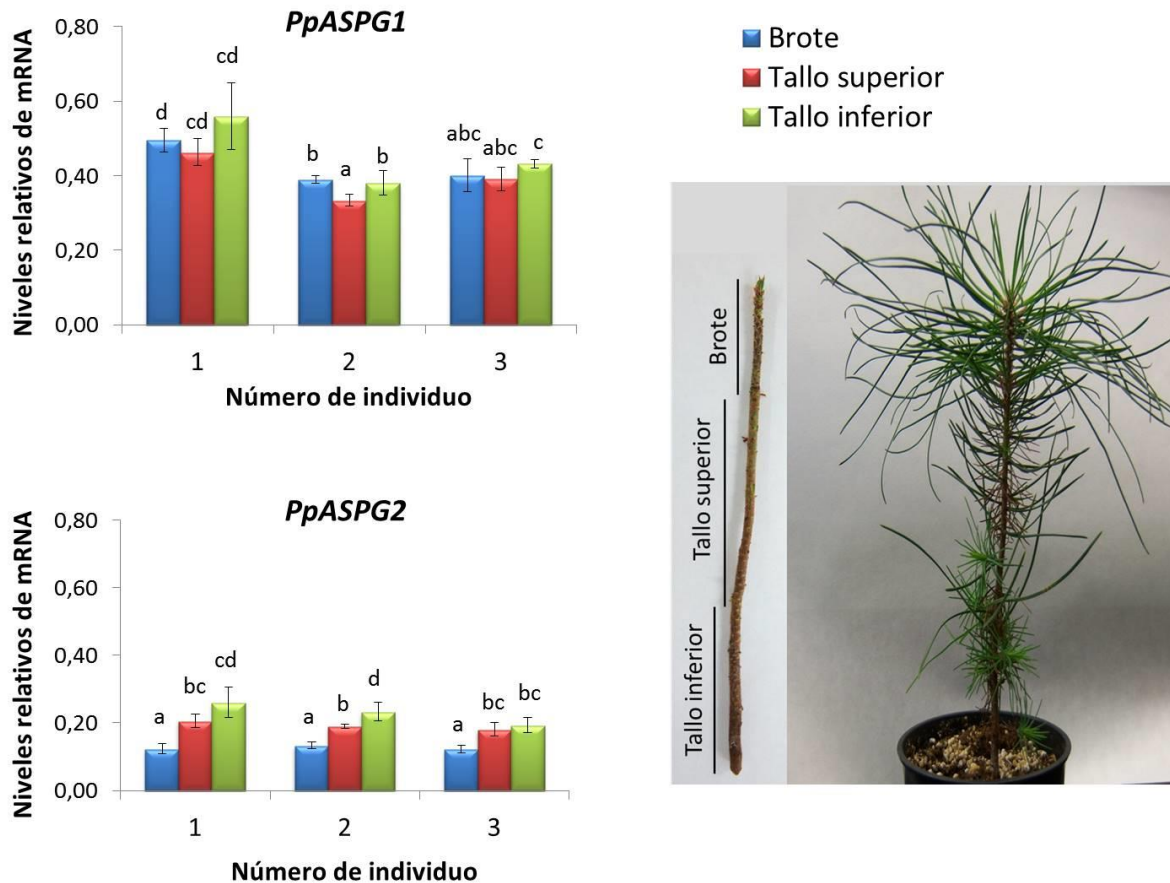


Figura 4.24 Niveles relativos de mRNA de los genes *PpASPG1* y *PpASPG2* a lo largo del eje vertical del tallo. Los niveles relativos de mRNA de los dos genes se determinaron por qRT-PCR en distintas secciones de tallos de plántulas con una edad de dos años. Los tallos de tres individuos se dividieron en tres partes como muestra la foto y se representan las medias de los niveles relativos de tres réplicas \pm SD. Los valores significativamente diferentes según el test *t* de Student ($P \leq 0,05$) están marcados con letras distintas.

Para evaluar la actividad del gen *PpASPG1* en individuos adultos se cuantificó los niveles de sus transcritos en muestras de árboles de 20 años de edad, tomadas en dos meses diferentes del año, marzo y mayo, cuando se espera una mayor actividad metabólica en el cambium vascular relacionado con la formación de la pared celular y el crecimiento secundario. La figura 4.25 muestra los niveles de mRNA relativos en las muestras recolectadas en marzo. En todos los individuos observamos que el xilema en diferenciación es el tejido

con el menor nivel de transcritos, mientras que en acículas, brotes y la muestra de floema y cambium observamos mayores niveles relativos de transcrito, aunque variable entre individuos.

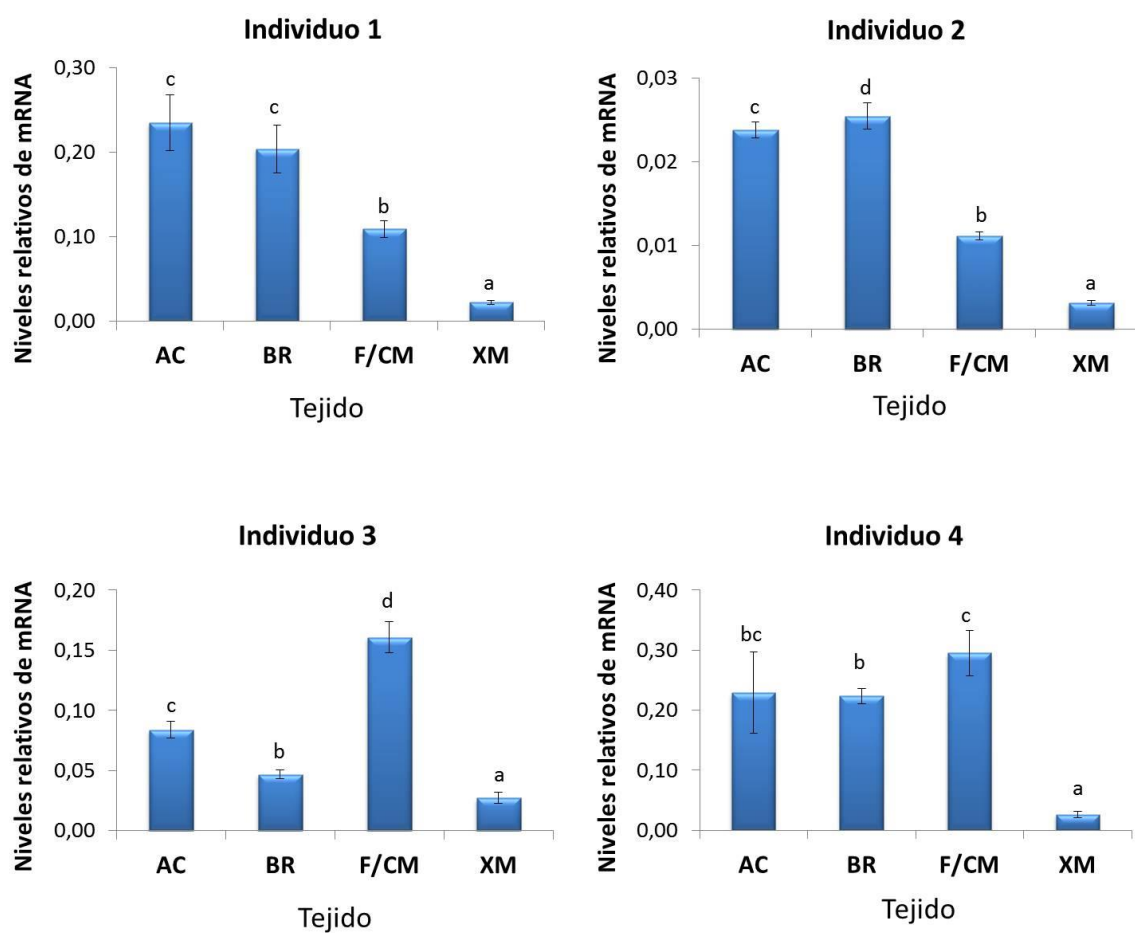


Figura 4.25 Niveles relativos de mRNA de *PpASPG1* en tejidos de árboles adultos muestreados en marzo. La abundancia relativa de los transcritos de *PpASPG1* se determinó en 4 árboles adultos en el mes de marzo del año 2008. Los tejidos analizados son acículas (AC), brotes (BR), floema maduro que contiene también la región del cambium vascular (F/CM) y xilema maduro (XM). Los valores representan tres réplicas técnicas \pm SD y los que tienen distintas letras son significativamente diferentes.

En la figura 4.26, que muestra los resultados de las muestras recogidas en el mes de mayo, se observa que el xilema maduro sigue siendo el tejido con la menor cantidad de transcritos, mientras que en el resto de las muestras analizadas se observa algo similar a lo descrito para el mes de marzo, presentando una mayor variación en los niveles entre individuos.

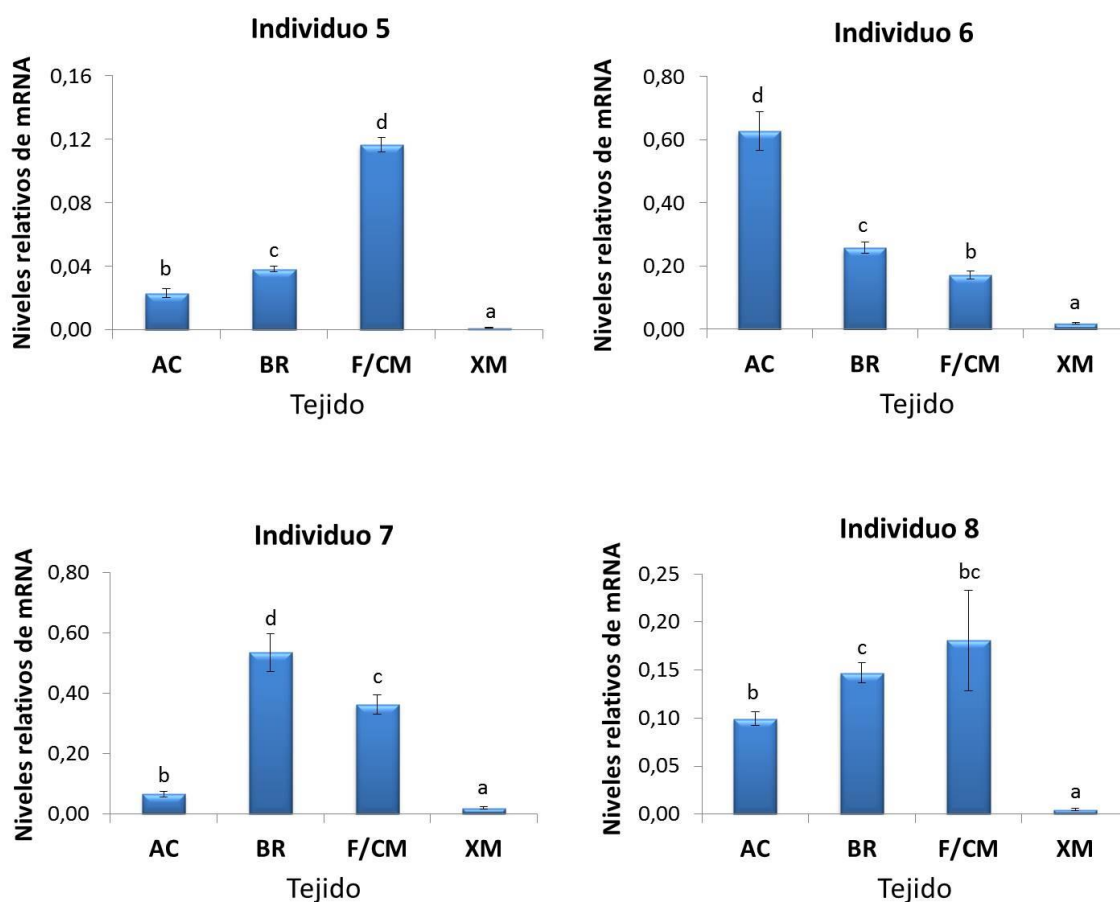


Figura 4.26 Niveles relativos de mRNA de *PpASPG1* en tejidos de árboles adultos muestreados durante el mes de mayo. La abundancia relativa de los transcritos de *PpASPG1* se determinó en 4 árboles adultos en el mes de mayo del año 2008, tal y como se ha descrito en la figura 4.25.

4.6.2. Respuesta de los genes *ASPG*, *δ-OAT* y *GS1b* a la disponibilidad de nitrógeno

Varios trabajos publicados han sugerido una relación entre la expresión de la asparraginasa y el estado de nitrógeno interno o el balance N/C en la planta (Scheible *et al.*, 2004; Gutiérrez *et al.*, 2007; Canales *et al.*, 2010). Con el propósito de entender mejor el papel de las asparraginasas en la adaptación de la plántula a situaciones con distinta disponibilidad de nitrógeno, analizamos los niveles relativos de los mRNA en diferentes condiciones nutricionales. Se usaron plántulas en el estadio E4, las cuales habían consumido completamente el megagametofito y la cubierta seminal o tegumento se había separado de los cotiledones. Las plántulas se dividieron en tres grupos, los cuales se regaron una vez por semana durante un total de 5 veces con una solución nutritiva sin ninguna fuente de nitrógeno o suplementada con sulfato de amonio a 3 mM o 10 mM. La abundancia relativa de transcritos fue determinada a corto plazo (24 horas después del primer riego, Figura 4.27) y a largo plazo (29 días después del primer riego, Figura 4.28). Se observaron cambios significativos para todos los genes y en casi todos los órganos entre los grupos que disponían de N con respecto al grupo sin N, tanto a corto plazo como a largo plazo. La cantidad de mRNA del gen *PpASPG1* aumentó considerablemente a corto plazo, especialmente en cotiledones donde llegó a alcanzar un nivel hasta 12 veces más elevado que en las plántulas sin disponibilidad de nitrógeno y apenas se encontraron diferencias entre los tratamientos con 3 y 10 mM de $(\text{NH}_4)_2\text{SO}_4$. Los otros dos genes de asparraginasa mostraron una respuesta contraria. La inducción observada en los niveles de mRNA de *PpASPG1* persistió a largo plazo en cotiledón y radícula. Para el gen *PpASPG2* se aprecia una mayor disminución de los niveles de transcrito en cotiledón y radícula en respuesta al tratamiento a largo plazo con 10 mM de $(\text{NH}_4)_2\text{SO}_4$ en comparación con el tratamiento a corto plazo. El efecto negativo de los tratamientos con ambas concentraciones de $(\text{NH}_4)_2\text{SO}_4$ sobre los niveles de transcrito del gen *PpASPG3* en cotiledones y del tratamiento con 10 mM de

(NH₄)₂SO₄ sobre los niveles de transcrito en hipocótilo y radícula se mantuvo tras 29 días de tratamiento.

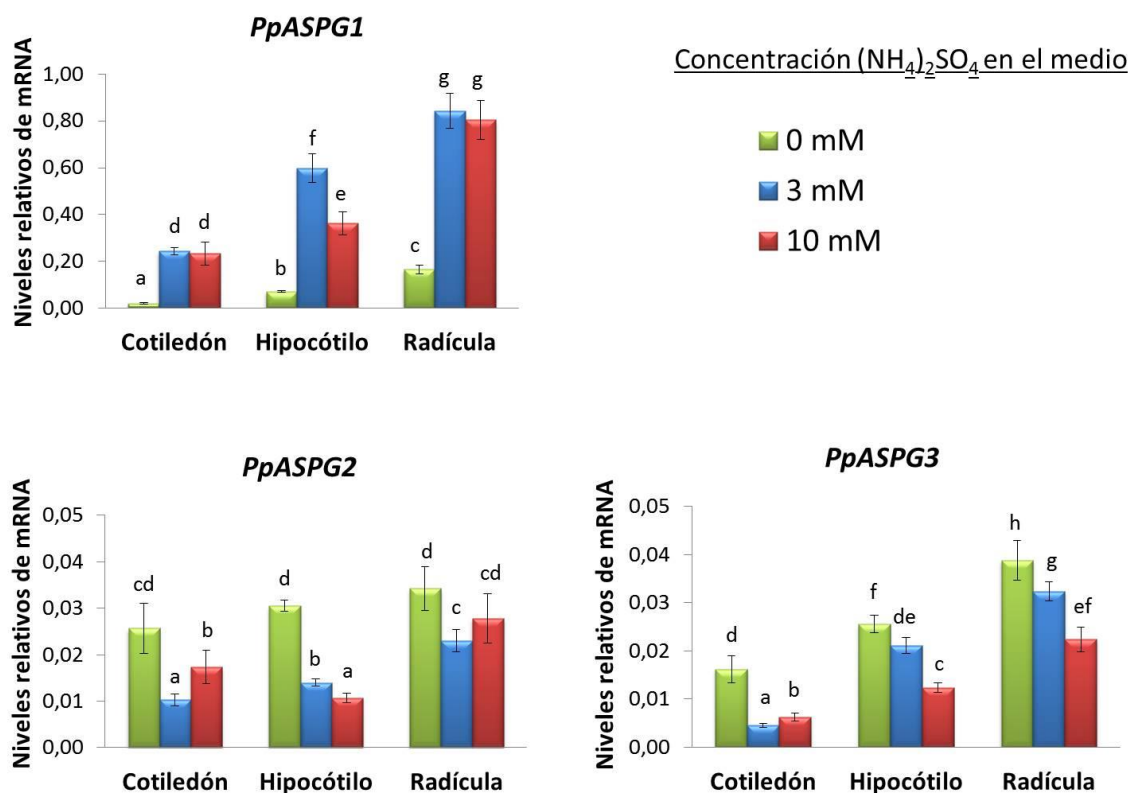


Figura 4.27 Abundancia relativa de transcritos de los tres genes de asparraginasa en respuesta a corto plazo a la disponibilidad de una fuente de nitrógeno. Tres grupos de plántulas del estadio de desarrollo E4 se regaron con distintas soluciones, un grupo con una solución nutritiva sin ninguna fuente de nitrógeno, el segundo grupo con la adición de 3 mM de sulfato de amonio y el tercer grupo con 10 mM de sulfato de amonio. Los niveles relativos de transcritos de los tres genes se determinaron mediante qRT-PCR a 24 horas tras los tratamientos. Los valores representan medios de 3 réplicas \pm SD. Los valores significativamente diferentes según el test *t* de Student ($P \leq 0,05$) son marcados con distintas letras.

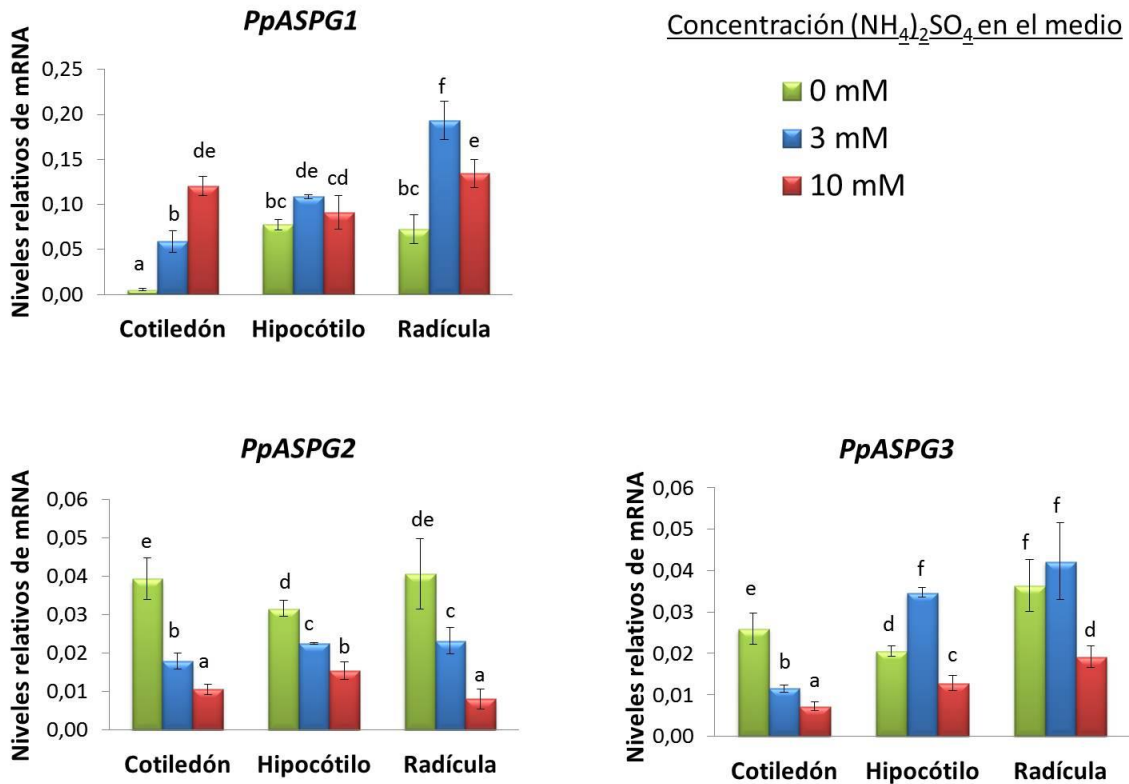


Figura 4.28 Abundancia relativa de transcritos de los tres genes de asparraginasa en respuesta a la disponibilidad de una fuente de nitrógeno a largo plazo. Los niveles relativos de mRNA se determinaron a largo plazo (29 días después del primer riego) y se presentan de la misma manera como en la figura 4.27.

Para obtener una visión más amplia del efecto de los tratamientos realizados sobre el metabolismo del nitrógeno en plántulas de pino, se analizaron también los niveles de mRNA de otros dos genes relacionados: la ornitina- δ -aminotransferasa (δ -OAT) y glutamina sintetasa (*GS1b*). En la figura 4.29 se observa una respuesta positiva a la presencia de una fuente de nitrógeno por parte del gen δ -OAT que resultó ser parecida a la observada para el gen *PpASPG1* con un aumento en la cantidad de transcritos y especialmente en cotiledones a largo plazo. Por el contrario, la única respuesta destacable del gen *GS1b* fue un aumento de sus niveles de transcritos en cotiledones tras ambos tratamientos con $(\text{NH}_4)_2\text{SO}_4$ a largo plazo.

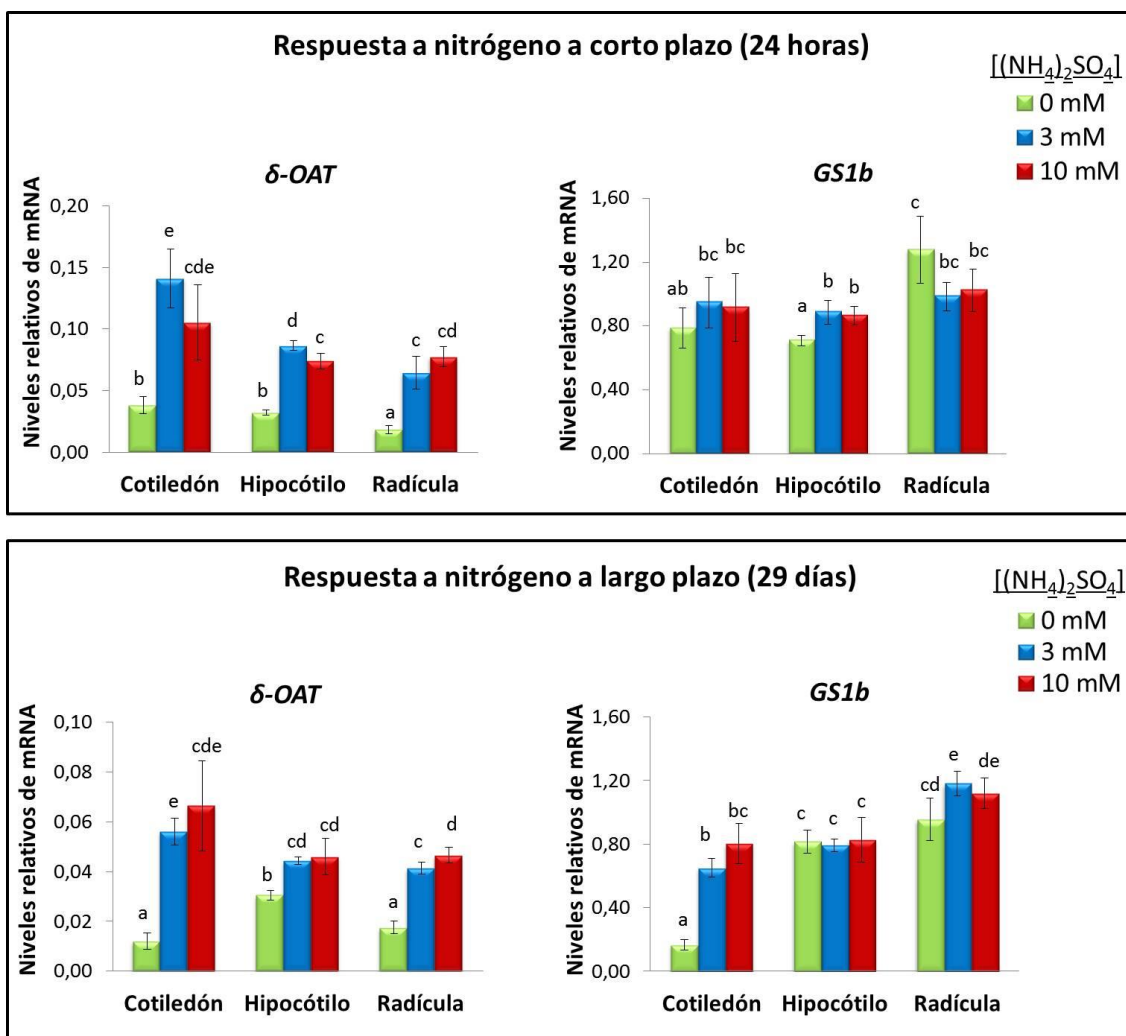


Figura 4.29 Niveles relativos de transcritos de los genes δ -ornitina aminotransferasa (δ -OAT) y glutamina sintetasa (*GS1b*) en respuesta a la disponibilidad de una fuente de nitrógeno. La abundancia relativa de transcritos de los dos genes se determinó de la misma manera como se ha descrito en la figura 4.27.

4.6.3 Efecto de factores de estrés sobre la expresión de *PpASPG1* y *PpASPG2* en plántulas

Análisis del transcriptoma en diversas especies de plantas, como soja, algodón, arabidopsis o *Taxus chinensis* han incluido a genes de asparraginasa entre aquellos cuya expresión varía en respuesta a factores de estrés o a hormonas relacionadas con la respuesta a estrés como el jasmonato (Cho *et al.*, 2007; Dubey *et al.*, 2013; Aceituno *et al.*, 2008; Li *et al.*, 2012). Para determinar si la ASPG de pino también tiene una función en los procesos de defensa o adaptación sometimos plántulas jóvenes a diferentes tratamientos. Conjuntos de plántulas en estadio E1 se dividieron en diferentes grupos experimentales y se determinaron los niveles relativos de mRNA de los genes *PpASPG1* y *PpASPG2* en diferentes tejidos y tiempos después de aplicarles un tratamiento simulando algunos factores de estrés o aplicando sobre la planta soluciones de ácido abscísico o ácido jasmónico. Se representa la respuesta de los genes como la razón entre la abundancia de transcritos en condiciones de estrés con respecto a plántulas control que no recibieron ningún tratamiento.

La figura 4.30 muestra los resultados observados para ambos genes a un tratamiento con una solución de jasmonato 100 μ M. Se observa que el nivel de transcritos del gen *PpASPG1* cambia rápidamente y varía dentro de la plántula espacial y temporalmente. La aplicación de la hormona produjo un aumento de mensajeros de 6 veces en cotiledones, pero una disminución en hipocótilos y radículas a tan sólo 2 horas tras aplicar la hormona. Con el paso del tiempo el efecto inductor se reduce en cotiledones, pero se traslada a los otros tejidos donde a las 24 horas se presenta un aumento hasta 7 veces en hipocótilos y de 2 veces en radículas. Para el gen *PpASPG2*, igualmente se observaron cambios en el nivel de mRNA en respuesta a la aplicación de jasmonato, pero de menor amplitud que los observados para el gen *PpASPG1*.

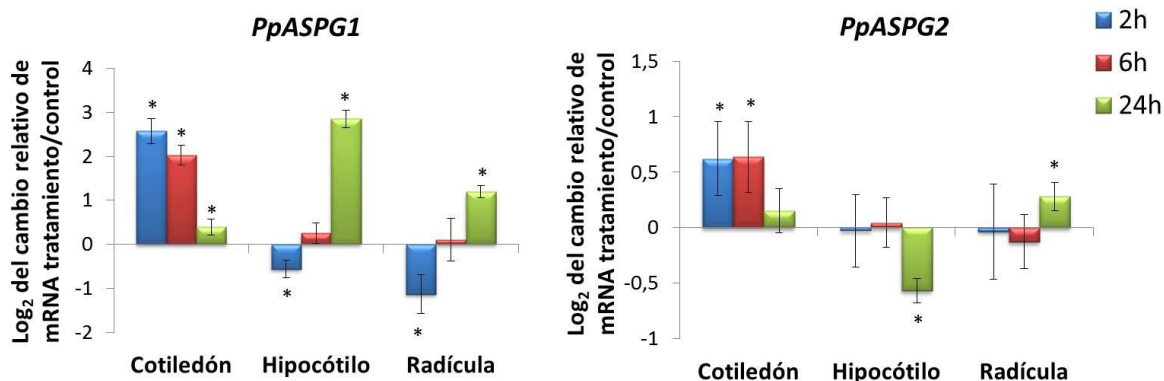


Figura 4.30 Abundancia relativa de mRNA de *PpASPG1* y *PpASPG2* en respuesta a jasmonato 100 μ M. Plántulas en estadio E1 se sometieron a un tratamiento con la hormona y los niveles relativos de mRNA de ambos genes se determinaron mediante qRT-PCR a 2h, 6h y 24h después de aplicar el tratamiento. Se representa la relación de transcritos en los grupos experimentales con respecto a las plantas control que no recibieron ningún tratamiento. Los valores significativamente diferentes a las muestras control según el test *t* de Student ($n = 3$) son marcados con * ($P \leq 0,05$). El cambio relativo se calculó sobre la media de las 3 réplicas técnicas de los grupos experimentales con respecto a la media del grupo control.

También una herida causó variaciones en los niveles de *PpASPG1* que dependieron del órgano y del tiempo tras la aplicación del estrés, pero que fueron más moderadas en general (Figura 4.31). La hormona ABA tuvo un efecto consistente de disminución de los niveles de transcrito de *PpASPG1* en todos los tejidos, que perduró al menos 24 horas. Aunque en el caso de radículas a 6 horas el cambio no fue estadísticamente significativo (Figura 4.32). El tratamiento de herida también causó una variación en los niveles de transcritos *PpASPG2* en los tres órganos tras dos horas de la aplicación del estrés, pero el cambio sólo se mantuvo tras 6 y 24 horas en cotiledones (Figura 4.31). Por otro lado, el tratamiento con la hormona ABA causó en *PpASPG2* una respuesta contraria a la del gen *PpASPG1*, resultando en un aumento del nivel de transcritos en todos los órganos tras 2 y 24 horas, pero no tras 6 (Figura 4.32).

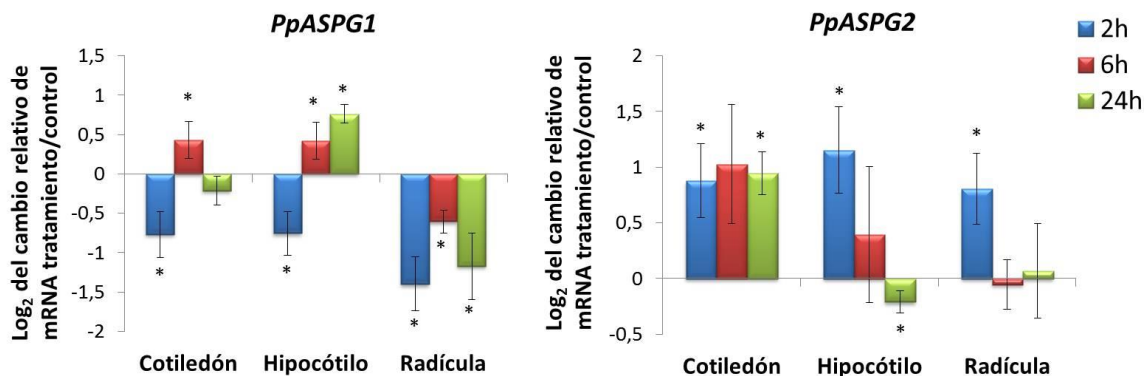


Figura 4.31 Niveles relativos de mRNA de *PpASPG1* y *PpASPG2* en respuesta a heridas provocadas en cotiledones e hipocótilos. Los niveles relativos de transcritos de ambos genes se determinaron en plántulas sometidas al tratamiento de estrés y se representan como se ha descrito en la figura 4.30.

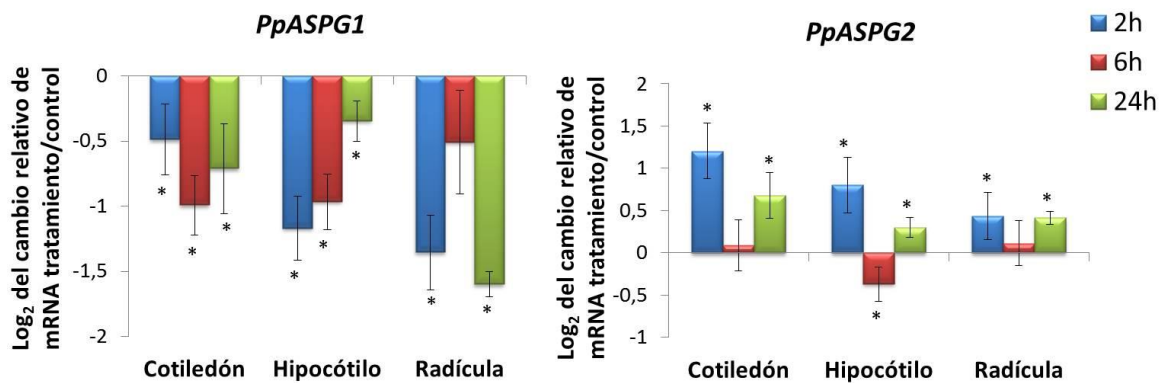


Figura 4.32 Abundancia relativa de transcritos de *PpASPG1* y *PpASPG2* en respuesta a ABA 50 μ M. Plántulas en estadio E1 se sometieron a un tratamiento con la hormona y los relativos niveles de mRNA se determinaron como se ha descrito para la figura 4.30.

Las respuestas del gen *PpASPG1* a frío y un exceso de sal se muestran en la figura 4.33. La presencia de una alta concentración de sal en el suelo causó una disminución moderada de la cantidad de transcritos en hipocótilo y radícula a las 48 horas tras el riego con NaCl 250 mM disuelto en agua. La aplicación de frío produjo un leve cambio significativo en hipocótilos a 48 horas donde se detectó un nivel relativo de 1,5 veces superior al de las plántulas controles. No detectamos alteraciones significativas en el nivel de mRNA de *PpASPG2* en respuesta a las condiciones frío y sal (datos no representados).

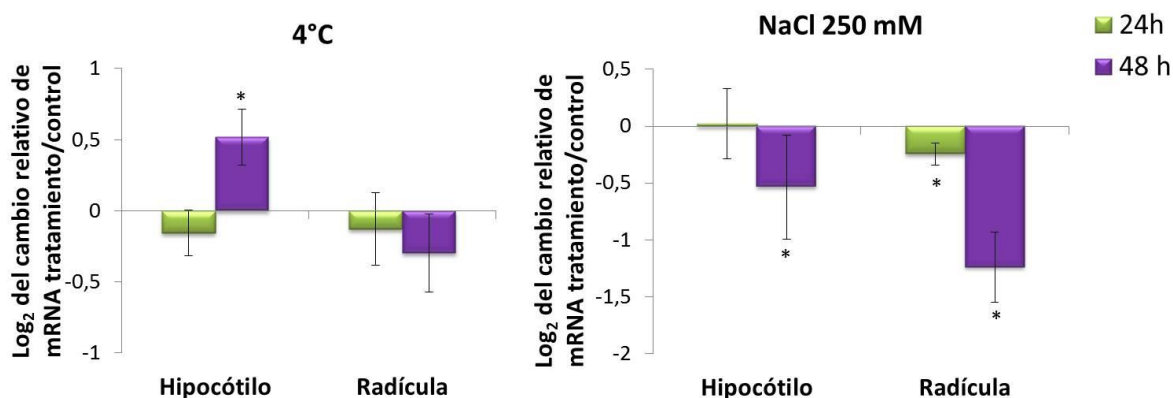


Figura 4.33 Abundancia relativa de transcritos de *PpASPG1* en respuesta a frío y sal.

Plántulas en estadio E1 se sometieron a los factores de estrés y los niveles relativos de mRNA se determinaron a 24h y 48h después del tratamiento en hipocótilos y radículas. Los valores se representan como se ha descrito para la figura 4.30.

Los resultados obtenidos sobre las características enzimáticas de las asparraginasas de *P. pinaster* y expresión de sus genes indican que *PpASPG1* es la enzima con una función relevante en la movilización de nitrógeno a partir de la molécula de reserva L-Asn. Por tanto, puesto que el interés de nuestro grupo de trabajo se centra en el metabolismo de nitrógeno y su relación con el desarrollo en coníferas, se procedió a una caracterización más detallada de esta proteína y su correspondiente gen, presentada en las secciones 4.3, 4.4 y 4.5.

5. DISCUSIÓN

5.1 Identificación y caracterización de asparraginasas de *Pinus pinaster*

5.1.1 El transcriptoma de pino marítimo revela la presencia de tres genes para ASPG

Basándonos en la especie *Pinus pinaster*, de la cual disponemos de un transcriptoma bastante completo (Canales *et al.*, 2014), se dedujo la existencia de, al menos, tres genes que codifican para proteínas relacionadas con asparraginasa en su genoma. El alineamiento y análisis de similitud de las asparraginasas identificadas en el transcriptoma de pino con las previamente caracterizadas en angiospermas agrupa PpASPG1 con el tipo dependiente de potasio y PpASPG2 y PpASPG3 con el grupo independiente del ion (Figura 4.1 a 4.3). Esta clasificación sugiere que el evento evolutivo de la separación entre los dos subgrupos de ASPG se produjo antes de la separación entre gimnospermas y angiospermas. En especies de briofitas también se han identificado más de un gen de asparraginasa, aunque en este caso no se distribuyen en los dos subgrupos (Figura 1.1). Comparando sus estructuras primarias con las ASPG de pino, todas las de musgos tienen el mayor porcentaje de identidad con PpASPG2 (63 % por término medio) y el menor con PpASPG1 (56 %). Con *A. thaliana* comparten un 61% de identidad con la ASPG independiente de potasio y 54 % con la forma dependiente de potasio. En el genoma de una especie de liquen (*S. moellendorffii*) sólo hay un gen de asparraginasa, igualmente con mayor parecido al grupo independiente de potasio. Aunque sería necesaria una comprobación experimental sobre la presencia de los dos grupos de asparraginasas en plantas no vasculares, el análisis basado en la conservación de la estructura primaria sugiere que la separación entre potasio dependiente e independiente implica una

especialización en función relacionada con el desarrollo de las plantas vasculares. No obstante, es difícil determinar el origen de los subgrupos debido a la baja cantidad de secuencias disponibles de especies no angiospermas. Por ejemplo, en las bases de datos de EST de helechos (*Ceratopteris richardii*) y de las gimnospermas *Ginkgo biloba* y *Cycas rumphii* no encontramos ninguna secuencia con una similitud considerable con las asparraginasas de plantas. Los genomas de las coníferas *Picea abies* y *Pinus taeda*, que han sido secuenciados, contienen tres secuencias que codifican para asparraginasas. Sin embargo, con la excepción de una ASPG de *P. taeda*, las secuencias son incompletas por lo cual la secuencia primaria de la proteína no se puede deducir completamente. Considerando la distancia evolutiva existente con especies angiospermas, como también refleja el árbol filogenético de la figura 4.2, podría haber diferencias en las características de las enzimas ortólogas. La ASPG, por su actividad de movilizar el N en tejidos sumideros, constituye un nexo esencial entre el metabolismo de los tejidos fotosintéticos y no fotosintéticos de la planta y entre el metabolismo primario y secundario. Por lo tanto, dado las peculiaridades de las leñosas perennes en general y de las coníferas en particular (Cánovas *et al.*, 1998, 2007; Cantón *et al.*, 2005) es de interés entender el papel específico de estas enzimas en el desarrollo y adaptación a cambios ambientales que, en el caso de las coníferas, se extienden en extensos periodos de tiempo.

5.1.2 Las tres ASPG de pino presentan doble actividad asparraginasas/isoaspartilo dipeptidasa

Con el objetivo de determinar la actividad bioquímica de las tres proteínas con similitud a asparraginasas identificadas a partir del transcriptoma de pino, acudimos a la expresión de las proteínas recombinantes en *N. benthamiana*. De esta forma, conseguimos cantidades suficientes de las proteínas activas, permitiendo su purificación y caracterización. La expresión en células de *E. coli* resultó en la recuperación de la proteína recombinante mayoritariamente en su

forma inmadura, como se había comprobado previamente expresando una asparraginasa de *Pinus sylvestris* en este sistema (Cañas *et al.*, 2007). El análisis de extractos de las hojas de *N. benthamiana* expresando las proteínas de pino demostró que las tres tienen la capacidad de catalizar la hidrólisis de asparragina, confirmando su anotación como asparraginasas (Figura 4.4). No obstante, observamos diferencias en el grado de procesamiento *in planta* con una maduración de toda la proteína presente en el caso de PpASPG1 y, por el contrario, la presencia de moléculas de precursor no procesado en el caso de PpASPG2 y PpASPG3. No existen referencias sobre la maduración *in planta* en otras especies, ya que la mayoría de los trabajos publicados tratan de proteínas recombinantes expresadas en *E. coli* y en este sistema heterólogo la eficiencia de maduración del precursor a la forma activa es inferior (Michalska and Jaskólski, 2006).

La caracterización de las tres ASPG de pino mostró que todas tienen una doble actividad L-asparraginasa/isoaspartilo dipeptidasa (Figura 4.5), como también se observó en *Arabidopsis thaliana* (Gabriel *et al.*, 2012). Para la ASPG dependiente de potasio de *A. thaliana* se determinó una mayor afinidad, pero una menor velocidad máxima, para dipéptidos de isoaspartilo que para asparragina. De tal forma, la eficiencia catalítica para asparragina fue 16 veces la observada para β -Asp-His. La ASPG de pino con más similitud a las del tipo dependiente de potasio (PpASPG1), a diferencia de su ortólogo en *arabidopsis*, tiene una K_m aparente para dipéptidos muy alta y no llegó a saturarse con el sustrato β -Asp-Ala a una concentración de 100 mM. Tiene una eficiencia casi 19 veces mayor para asparragina como sustrato que para β -Asp-Ala (Figura 4.6), por lo que su función primaria consiste en la movilización del nitrógeno contenido en el grupo amino del aminoácido. El valor de su K_m para la asparragina de 4,58 mM se sitúa dentro de los márgenes determinados para sus ortólogos en angiospermas con valores de 6,83 mM (Gabriel *et al.*, 2012) en *A. thaliana* y 7,2 mM en *L. japonicus* (Credali *et al.*, 2011).

Aunque PpASPG2 y PpASPG3 se clasifican en el grupo de ASPG independiente de potasio, en base a su estructura primaria, sólo PpASPG3 presentó una clara

preferencia por el dipéptido de isoaspartilo como sustrato. Con una K_m de 1,04 mM para β -Asp-Ala y una V_{max} de 2,03 U/mg, cataliza la hidrólisis de este sustrato con una eficiencia casi 22 veces mayor que la mostrada para asparragina, para la cual tiene una K_m de 6,12 mM y una V_{max} de 0,53 U/mg, asociando su función a la degradación de estos dipéptidos aberrantes. En *Lotus japonicus* la ASPG independiente de potasio tiene una K_m de 30 mM para asparragina y 0,54 mM para β -Asp-His (Credali *et al.*, 2011) y su ortólogo en *Lupinus luteus* presenta una K_m de 4,8 mM para asparragina y 0,14 mM para β -Asp-Leu (Borek *et al.*, 2004), datos comparables con nuestros resultados para PpASPG3. Por otro lado, aunque no pudo determinarse los valores de K_m para PpASPG2, se observó que la enzima cataliza las reacciones de ambos tipos de sustrato con una velocidad similar, lo que la diferencia de las otras dos ASPG de pino identificadas. No obstante, podría ocurrir que la enzima presentara mayor afinidad y/o V_{max} para algún dipéptido de isoaspartilo no ensayado en comparación a L-Asn como sustrato. Considerando el conjunto de trabajos publicados sobre las ASPG de plantas, existe gran variabilidad en las características bioquímicas dentro de este grupo de enzimas, entre especies y entre las diferentes isoformas de la misma especie, como también observamos en *P. pinaster*. Podría reflejar los distintos requerimientos de esta actividad en base a las características propias de la planta respecto al uso de asparragina y generación de dipéptidos de isoaspartilo, pero también podrían ser consecuencias de diferencias en los ensayos específicos y los sustratos ensayados por los diferentes autores.

El análisis estructural de una ASPG dependiente de potasio de *Phaseolus vulgaris* mediante cristalografía permitió identificar los residuos relacionados con la especificidad de sustrato (Bejger *et al.*, 2014). Además del residuo de treonina nucleófilo y los residuos de arginina y glutamato del interruptor catalítico, otros cuatro aminoácidos localizados en la subunidad β que también forman interacciones con el producto aspartato en el sitio activo (en color rosado en la figura 4.1) se conservan en todas las ASPG incluidas en el alineamiento, tanto las del grupo dependiente de potasio como las independientes. Los residuos que estructuralmente se localizan próximos a la

cadena lateral del sustrato (en color azul en la figura 4.1) parecen ser los determinantes de la preferencia de sustrato. Un motivo GAG formado por Gly11-Ala12-Gly13 en PvAspG1, supuestamente importante para permitir el anclaje de sustratos de isoaspartilo (Michalska *et al.*, 2006), también se conserva entre todas las secuencias con la excepción de PpASPG3 donde el motivo cambia a GAR. Curiosamente PpASPG3 presenta mucha afinidad para β -Asp-Ala luego este cambio parece no afectar negativamente a la actividad de la enzima en coníferas. Los 9 aminoácidos que siguen a este motivo, igualmente identificados como importantes para la especificidad de sustrato, varían considerablemente entre las diferentes enzimas. Uno de estos aminoácidos, Glu20 de PpASPG1, está conservado entre las ASPG de las cuales se determinó experimentalmente la doble actividad, incluyendo las ASPG dependiente e independiente de potasio de *A. thaliana* (Gabriel *et al.*, 2012). La ASPG de *L. japonicus* del grupo dependiente de potasio, para la cual no se registró una actividad dipeptidasa, presenta un residuo de glutamina en esta posición (Gln21 en LjNSE1), lo que sugiere que el aminoácido ácido es importante para esta actividad. Dentro de la subunidad β se identificó otra secuencia relacionada con la especificidad de sustrato (Bejger *et al.*, 2014), donde PpASPG1 curiosamente presenta dos aminoácidos invertidos en su posición: Glu334-Asp335 con respecto a las asparraginasas dependientes de potasio de angiospermas (Asp285-Glu286 en PvAspG1 y LjNSE1). Según Credali *et al.* (2011) estos dos aminoácidos podrían contribuir a la coordinación con el ion potasio, pero los estudios estructurales de PvAspG1 no respaldan esta suposición (Bejger *et al.*, 2014). El resto de los aminoácidos importantes para la especificidad de sustrato se conservan bien entre las ASPG dependientes de potasio, pero en el otro subgrupo las secuencias de las ASPG de pino difieren más de sus ortólogos en angiospermas, lo que puede indicar una mayor variabilidad en la preferencia de sustratos dentro de las ASPG independientes de potasio.

5.2 Características particulares de la proteína PpASPG1 y función de la región variable

5.2.1 Las asparraginasas de coníferas del tipo dependiente de K⁺ presentan la región variable de mayor extensión entre las ASPG caracterizadas

Las asparraginasas de plantas, como miembros de la familia de NTN-hidrolasas, se expresan en forma de precursor que sufre una activación proteolítica que genera el residuo nucleófilo en posición amino-terminal esencial para la actividad. El análisis mediante espectrometría de masas de la ASPG independiente de potasio de *A. thaliana* expresada en células de *E. coli* mostró que la activación era resultado de una única reacción de hidrólisis del enlace peptídico previo al residuo nucleófilo y no de la eliminación de un péptido interno (Hejazi *et al.*, 2002). Sin embargo, en el caso de la ASPG recombinante de *E. coli* EcAIII, tras el proceso de maduración, eventos de proteólisis *in vitro* adicionales resultaron en un acortamiento en la secuencia de ambas subunidades sin afectar a la actividad catalítica (Borek *et al.*, 2004). La caracterización de la estructura de la proteína independiente de potasio de *Lupinus luteus* recombinante cristalizada sugirió que la región C-terminal de la subunidad α presenta una secuencia no estructurada (Michalska *et al.*, 2006). Esta secuencia, que se corresponde con la RV, difiere tanto en longitud como en secuencia entre especies e isoformas de la enzima. En el caso de las ASPG dependientes de potasio de coníferas la RV es notablemente más extensa que en las proteínas parálogas y ortólogas y, además, la secuencia no presenta similitud con ninguna otra proteína eucariota que no sea de coníferas, con la excepción de los últimos 13 residuos que preceden al residuo nucleófilo Thr245 (Figura 4.1). El análisis de la composición de la proteína madura recombinante mediante espectrometría de masas mostró que la RV completa de PpASGP1 forma parte de la enzima madura, hasta el residuo aminoacídico Glu244 (Figura 4.8). En *L. japonicus* también se demostró que la secuencia completa

de la RV es mantenida en la proteína madura (Credali *et al.*, 2011), pero en este caso la región contiene sólo 28 residuos mientras que en PpASPG1 presenta un total de 77 aminoácidos. Algunas asparraginasas de briofitas (PhPaten2 y PhPaten3 en el árbol filogenético de la figura 1.1) también tienen una región variable de alrededor de 60 residuos, pero la secuencia no tiene similitud con la de PpASPG1, ni siquiera en características químicas de los residuos de aminoácidos que la componen.

Dado que se mantiene en la proteína madura de coníferas, la RV no debe interferir negativamente con la actividad básica de la enzima, ya que esta presenta características cinéticas similares a las de angiospermas que carecen de una RV tan extensa. Esto no impide que la presencia de esta extensa región pueda tener efectos sobre el proceso de maduración o aspectos específicos de la función enzimática.

5.2.2 Modificaciones en la región variable afectan a la maduración del precursor *in planta* e *in vitro*

El procesamiento del precursor de las ASPG de plantas ha sido abordado por varios autores y este proceso de activación en especies angiospermas se ha caracterizado como una maduración autoproteolítica (Borek and Jaskólski, 2001; Bruneau *et al.*, 2006; Michalska *et al.*, 2006; Credali *et al.*, 2011; Gabriel *et al.*, 2012; Bejger *et al.*, 2014). Sin embargo, una ASPG recombinante de *Pinus sylvestris* homóloga a PpASPG1 expresada en células de *E. coli* no presentaba la capacidad de autoprocesamiento espontáneo *in vitro* (Cañas *et al.*, 2007). La imposibilidad de modelizar la estructura de la RV en la ASPG dependiente de potasio de *P. vulgaris*, debido a una falta de densidad electrónica en los análisis de difracción de rayos X, limitó la capacidad de poder asignarle funciones específicas (Bejger *et al.*, 2014). Para determinar la implicación de la RV de PpASPG1 en su maduración comparamos la eficiencia del procesamiento *in planta* entre diferentes variantes de la proteína

recombinante en las cuales se eliminaron diferentes fragmentos de esta región. Las secuencias eliminadas fueron escogidas de tal manera que el residuo precedente al nucleófilo Thr245 fuera siempre glutamato. Ninguna de las mutaciones en la proteína recombinante causó una supresión completa del procesamiento del precursor. Sin embargo, las alteraciones de la RV afectaron negativamente a la eficiencia del procesamiento del precursor. Por lo que, nuestros resultados indican que alteraciones de la RV afectan negativamente a la eficiencia del procesamiento del precursor, ya que sólo en el caso de la proteína con esta región intacta se observó el procesamiento del 100% de la proteína producida (Figura 4.10). En *A. thaliana*, la ASPG dependiente de potasio toleró la eliminación de pequeños fragmentos en la RV y también tuvo un impacto sobre el procesamiento y la actividad catalítica (Gabriel *et al.*, 2012). Por otro lado, los resultados del análisis de la estructura de la enzima de *P. vulgaris* previamente comentados sugiere que la RV de las ASPG en especies angiospermas no es estructurada (Bejger *et al.*, 2014) y una cierta flexibilidad en la zona que precede al nucleófilo es necesaria para permitir el autopprocesamiento (Nomme *et al.*, 2012), pero no disponemos de información sobre la estructura de las ASPG en gimnospermas.

Generalmente se observa que las asparraginasas recombinantes expresadas en células de *E. coli* tienen la capacidad de autopprocesarse *in vitro* (Borek *et al.*, 2004; Michalska and Jaskólski, 2006; Cantor *et al.*, 2009; Nomme *et al.*, 2012). La expresión de PpASPG1 y de PpASPG1-D3 en este sistema heterólogo resultó en una alta producción de proteínas recombinantes, pero mayoritariamente en forma de precursor no procesado. Cuando se comparó la capacidad de autopprocesamiento de ambas proteínas *in vitro* solamente la proteína PpASPG1-D3 mostró procesamiento de una pequeña parte de la proteína total, sugiriendo que la RV disminuye la capacidad de autoproteólisis *in vitro* de la proteína recombinante (Figura 4.12). La ASPG recombinante de *Pinus sylvestris* tampoco presentaba la capacidad de autopprocesamiento *in vitro* aunque ésta fue estimulada mediante la adición de extractos de radículas de pino (Cañas *et al.*, 2007). Estos resultados sugieren que la RV es requerida para optimizar el proceso de la maduración del precursor *in planta* y podría

estar implicada en un mecanismo que requiera de algún factor o de ciertas condiciones específicas para activarse.

5.2.3 Papel de la región variable en las características bioquímicas de PpASPG1

5.2.3.1 Efecto de la eliminación de la región variable sobre los parámetros cinéticos de la enzima

Para evaluar el efecto de la ausencia de la RV de PpASPG1 en sus características enzimáticas comparamos las mismas en PpASPG1 y PpASPG1-D3 (Figura 4.13). La V_{\max} de PpASPG1-D3 para la L-asparagina se redujo a aproximadamente un cuarto de la enzima de secuencia completa, probablemente debido a una gran proporción de proteína que se encuentra en forma de precursor inactivo en la muestra. PpASPG1 hidroliza asparagina con una eficiencia casi 19 veces mayor a la que muestra para β -Asp-Ala, pero sorprendentemente la eliminación de la RV produce una enzima con casi idénticos parámetros cinéticos para ambos sustratos. Los resultados sugieren que la RV reduce considerablemente la capacidad de hidrólisis de β -Asp-Ala en favor del sustrato más pequeño. Aunque esta secuencia variable no está resuelta en la estructura de *P. vulgaris*, de dicha estructura se deduce que queda localizada en la entrada del centro activo (Figura 1.2). Por tanto, en el caso de PpASPG1 podría dificultar la entrada o posicionamiento de dipéptidos en el sitio activo, lo que determinaría su menor afinidad por dipéptidos de isoaspartilo. Para las ASPG de *A. thaliana* también se observó que la RV de la enzima dependiente de potasio influye en la preferencia de sustrato, pero de forma distinta (Gabriel *et al.*, 2012). Proteínas quimeras del tipo dependiente de potasio con la secuencia variable sustituida por la de su parólogo independiente de potasio mostraron una capacidad autoproteolítica y una V_{\max} considerablemente reducida, pero con sólo un pequeño impacto sobre la K_m

para los sustratos asparragina o dipéptidos de isoaspartilo. La sustitución recíproca resultó en una enzima con una afinidad aumentada para L-asparragina, con una reducción de 320 veces en el valor de K_m y una actividad catalítica para asparragina 60 veces más eficiente y no afectó al procesamiento autocatalítico.

5.2.3.2 La eliminación de la región variable disminuye la estabilidad de PpASPG1

Con el objetivo de estudiar si la RV influye en otras características de la enzima, analizamos la actividad asparraginasa en medios tamponados a diferentes pH. Las dos formas de la proteína PpASPG1 y PpASPG1-D3 mostraron una mayor actividad catalítica a pH 8 (Figura 4.14). Sin embargo, la eliminación de la misma resulta en la inactivación de la enzima a pH 6, pH en el que PpASPG1 mantiene un 70% de su actividad. Cabe la posibilidad que His187 en particular, la única histidina localizada en la RV, con un valor pK_a de 6,04 para el grupo ionizable de su cadena lateral, contribuya a la actividad de la enzima o a estabilizarla a este pH.

Aparte de diferencias en la afinidad y especificidad para el sustrato entre PpASPG1 y PpASPG1-D3, las 2 proteínas recombinantes presentaban diferencias en su temperatura óptima y energía de activación (Figura 4.15). PpASPG1 resultó más activa a 50°C mientras que PpASPG1-D3 ya sufre inactivación térmica a 45°C. Hay poca información disponible sobre la termoestabilidad de ASPG de plantas. La temperatura óptima determinada para una ASPG de *Vigna unguiculata* fue de 40°C (Ali, 2009) y la asparraginasa independiente de potasio de *Lupinus luteus* pierde actividad enzimática a 60°C (Borek *et al.*, 2004). Asparraginasas presentes en muestras de suelo, de origen desconocidas, tenían mayor actividad a 60°C y su energía de activación tenía un valor entre 20,2 y 34,1 kJ/mol (Frankenberger and Tabatabai, 1991). La ASPG es una de las enzimas que se encuentra libre en el suelo y su

actividad se usa, entre otras enzimas, para evaluar la calidad del suelo, pero podría ser de origen bacteriano (Acosta-Martinez *et al.*, 2014). La energía de activación calculada para una ASPG de *Pisum sativum* fue de 26,2 kJ/mol, un valor próximo a nuestro resultado para PpASPG1 de 27,98 kJ/mol. La mayor energía de activación determinada para PpASPG1-D3 de 34,24 kJ/mol sugiere que la RV contribuye a reducir de alguna manera la barrera de energía, resultando en una mayor velocidad de reacción.

La menor estabilidad de PpASPG1-D3 (Figura 4.16) sugiere también que la RV contribuye a la estabilidad de la proteína frente a la inactivación térmica de la enzima de alguna manera. Una estrategia similar aplicada sobre una Fd-GOGAT de *Synechocystis sp.*, de la cual se eliminó una secuencia variable de 27 aminoácidos, produjo una proteína activa con una V_{\max} ligeramente mayor pero menor afinidad para glutamina, que además era considerablemente menos estable que la proteína nativa de secuencia completa (Tripathy *et al.*, 2015). Se ha demostrado en una arquea termófila que secuencias no estructuradas conectoras entre dominios estructurados pueden contribuir a la estabilidad de la proteína y además pueden actuar como chaperonas para conseguir un correcto plegamiento (Tomar *et al.*, 2013). Posiblemente los efectos de la RV sobre las características bioquímicas de PpASPG1 se deben meramente al efecto de su presencia sobre la estructura cuaternaria de la proteína, pero debido a la poca información disponible sobre la estructura secundaria de esta secuencia en particular, es difícil postular sobre una posible función.

5.2.4 Efecto del ion potasio sobre la estabilidad y la actividad catalítica de PpASPG1

Las ASPG de plantas fueron clasificadas por Sodek *et al.* (1980) como dependientes o independientes de potasio. Posteriormente, se describieron asparraginasas dependientes de potasio en estudios realizados en otras

especies como *Arabidopsis thaliana* (Bruneau *et al.*, 2006), *Lotus japonicus* (Credali *et al.*, 2011) y *Phaseolus vulgaris* (Bejger *et al.*, 2014). En estas especies las asparraginasas de este subgrupo son entre 10 y 11 veces más eficientes en la hidrólisis de asparragina cuando se incluye KCl 50 mM en el medio de ensayo. Según los resultados de Bejger *et al.* (2014) el catión monovalente se une en un "bucle de activación" formado por unos residuos localizados en una secuencia conservada en la subunidad α (Figura 4.1). En nuestros experimentos tanto PpASPG1 como PpASPG1-D3 presentaba mayor actividad en presencia de KCl 50 mM, aunque la ausencia del catión sólo redujo la velocidad de reacción un 25% en el caso de PpASPG1-D3 y 35 % para la proteína de secuencia completa (Figura 4.17a). El papel del potasio en la estimulación de la actividad de la enzima no tiene la misma importancia que la descrita para especies angiospermas, en las cuales la proteína suele perder un 90 % de su actividad en ausencia del catión. Las ASPG dependientes de potasio de angiospermas comparten el motivo VMDKSPHS en el bucle de activación, pero en coníferas la secuencia correspondiente (VMERSPHV) tiene un total de tres residuos diferentes (Figura 4.1). Probablemente el cambio de Asp a Glu112 y Lys a Arg113 en PpASPG1 no sea responsable de la gran diferencia en el efecto del potasio ya que son aminoácidos con cadenas laterales similares, aunque en ambos casos el aminoácido de la ASPG de pino es la forma con cadena lateral más voluminosa. Sin embargo, el último residuo del motivo, serina en angiospermas, es aparentemente muy importante para el correcto posicionamiento del catión (Bejger *et al.*, 2014) y tampoco está conservada en PpASPG1 donde cambia a un aminoácido hidrofóbico Val117. Los residuos del bucle de activación no influyen directamente en el mecanismo catalítico, pero la coordinación con el ion metálico reordena el bucle tal que un cambio estructural es transmitido al sitio activo a través de tres residuos que forman un "interruptor catalítico" (His117, Arg224 y Glu250 en PvAspG1 en Figura 4.1) y que están conservados entre todas las secuencias de ASPG, tanto dependiente como independiente de potasio. En las ASPG independientes de potasio, que aparentemente tienen el interruptor permanentemente en posición "activa", el último aminoácido en la secuencia del bucle de activación es un

aminoácido hidrofóbico al igual que PpASPG1. Por otro lado, las ASPG dependientes de potasio de angiospermas tienen un residuo nucleófilo en esta posición y requieren de la coordinación con potasio para activar el interruptor, aunque sin el ion la enzima mantiene un nivel basal de alrededor de un 10 % de su capacidad catalítica máxima. Luego, en base exclusivamente a la secuencia del bucle de activación, PpASPG1, se clasificaría en el grupo de ASPG independiente de potasio, aunque el resultado en los ensayos en ausencia o presencia de K^+ sugiere que PpASPG1 presenta una dependencia limitada al ion. No obstante, la mayor actividad registrada para PpASPG1 en presencia de potasio puede deberse sólo a un aumento en la estabilidad de la proteína y no ser un efecto directo de activación. En una ASPG de *Pisum sativum* se demostró que el potasio contribuye a mantener la integridad funcional de la enzima (Sodek *et al.*, 1980). Nuestros experimentos confirman esta propiedad para el ion metálico y la protección contra la inactivación térmica era evidente para PpASPG1, pero aún más para PpASPG1-D3 (Figura 4.17b) que ya se había mostrado menos estable debido a la carencia de la RV (Figura 4.16). No obstante, nuestros resultados indican que la RV no influye en el efecto protector del ion ni en la capacidad de unión entre el potasio y la enzima.

La conservación de la RV en PpASPG1 es importante para la enzima en el pino ya que afecta a diversas características de la enzima: aumenta la estabilidad, favorece la actividad L-asparaginasa sobre la isoaspartilo dipeptidasa, reduce la energía de activación de la enzima y además podría permitir un control postraducciona sobre el proceso de maduración *in planta*.

5.3 Características de la región promotora de *PpASPG1*

Con el propósito de entender mejor la regulación transcripcional de *PpASPG1* aislamos la región promotora del gen (Figura 4.18) y la comparamos con secuencias promotoras disponibles en las bases de datos de algunas otras ASPG de especies angiospermas y gimnospermas (Figura 4.19). Observamos

una estricta conservación de una secuencia proximal, de 471 nucleótidos en el caso de *P. pinaster*, entre las ASPG potasio dependiente de las tres especies de pino (93 a 97 % de identidad) y un alto parecido con la de *Picea abies* (71 % de identidad). Curiosamente la similitud en secuencia se pierde totalmente en la región distal del promotor, no sólo entre especies sino también entre dos aislados independientes que amplificamos a partir del genoma de *P. pinaster*. Por lo que asumimos que los principales motivos para el control transcripcional deberían localizarse estrictamente en la secuencia proximal. En esta región proximal también se encuentran elementos *cis* relacionados con la regulación por hormonas y factores de estrés, para los cuales también determinamos experimentalmente una respuesta en la abundancia de transcritos del gen. No obstante, los resultados del análisis funcional *in planta* muestran una diferencia importante en los niveles de expresión del gen delator promovidos por las dos versiones de promotor que aislamos de *PpASPG1* (Figura 4.20b). Curiosamente las secuencias promotoras de *Pinus sylvestris* y *Picea abies* apenas presentan elementos *cis* relacionados con el control por hormonas o factores de estrés en su región distal (Figura 4.19). Una hipótesis a considerar podría ser que la inserción de un transposón y/o eventos de recombinación introdujeron posibles elementos de regulación transcripcional que contribuirían a modular la respuesta del gen a diversos factores. En concreto, que la secuencia de 231 nucleótidos que codifica para un fragmento de MDH se insertara en el promotor original dando lugar a la variante alélica de la que se derivó el fragmento Pin-GW2 aislado (Figura 4.18), como resultado de un evento de transposición. En el genoma de *Pinus taeda* este fragmento también aparece un número de veces, como por ejemplo en la secuencia con referencia lcl|scaffold209167 (congenie.org) dónde se repite dos veces con una secuencia de 673 pares de bases en medio. Resulta interesante que en el mismo scaffold, siguiendo a 5 kilobases del segundo fragmento MDH, se encuentra una secuencia de 656 nucleótidos muy parecida (90 % de identidad) a una secuencia en la región distal tipo Pin-GW1 de *PpASPG1*. Este scaffold, no obstante, no contiene ninguna secuencia parecida al promotor proximal o secuencia codificante para ASPG, aunque continúa un total de 17 y 61

kilobases hacia 5' y 3' de la secuencia tipo Pin-GW1 respectivamente. El hecho de que secuencias de localización cercana en el genoma de *P. taeda* se distribuyen entre dos versiones de promotor distal para *PpASGP1* en *P. pinaster* respalda la hipótesis de que eventos de recombinación y/o transposición se han producido en el promotor de esta especie. Retroelementos y derivados son ubicuos y abundantes en los genomas de plantas (Friesen *et al.*, 2001), pero especialmente en pinos donde el genoma gigante actual es, en parte, resultado de la actividad de retrotransposones y las regiones codificantes representan sólo una fracción minúscula del genoma (Neves *et al.*, 2013). Aparte de diferencias en el tamaño del genoma hay evidencias de una gran variación intraespecífica correlacionada con diferencias en hábitat y morfología e incluso diferentes poblaciones pueden diferir en su volumen nuclear (Murray, 1998).

Para estimar la variabilidad en la región reguladora del gen *PpASPG1* entre diferentes poblaciones de pinos se procedió a amplificar mediante PCR y secuenciar las secuencias promotoras en 11 poblaciones diferentes de *P. pinaster*. Aunque no se consiguió amplificar los fragmentos completos previamente amplificados por Genome Walker, si se consiguieron amplificados de la región proximal para todas las poblaciones. La comparación de las secuencias de los fragmentos aislados permitió identificar un número de SNPs, que conllevan diferencias en elementos *cis* relacionados con el control de transcripción por fitohormonas (Figura 4.21). Dentro de la secuencia total de 661 pb comparada entre las 11 poblaciones hubo un total de 8 SNPs, un valor medio de un SNP por 83 pb. La tasa de SNPs en árboles es alta, con una frecuencia media de un SNP por 100 pb (Neale and Kremer, 2011). No obstante, la frecuencia de SNPs depende de la región genómica, como observado en *Pinus monticola*, donde se identificó en promedio un SNP por 285 pb en secuencias de genes altamente conservados y un SNP por 81 pb en genes codificantes para proteínas con dominios de unión a nucleótidos y repeticiones ricos en Leu (Liu *et al.*, 2014). En la secuencia codificante para *PpASPG1* en la población de Oria detectamos un cambio no sinónimo (Gln23 → Lys23), también compartido con la secuencia aislada tipo Pin-GW1, que

además coincide en esta posición con la secuencia de ASPG potasio independiente de *Lupinus luteus* que figura en el alineamiento de la Figura 4.1. Las variantes alélicas identificadas podrían tener un impacto sobre la capacidad de adaptación de las distintas procedencias de *Pinus pinaster*. En este sentido, *PpASPG1* podría resultar un gen de interés en la explotación de la variabilidad natural para la selección de individuos con características optimizadas de respuesta a factores ambientales en vista a la producción forestal.

5.4. Análisis de los niveles de mRNA para asparraginasas en el pino

5.4.1 Función de las ASPG durante el desarrollo de pino

El uso de la asparragina como molécula de reserva y transporte de nitrógeno difiere entre especies, tejidos de una misma planta y de forma temporal (Lea *et al.*, 2007). Tras la germinación de las semillas de pino parte del N almacenado en el megagametofito que está en contacto con los cotiledones es realojado en forma de asparragina como molécula de reserva en hipocótilos, debido a la actividad asparragina sintetasa (Cañas *et al.*, 2006). En estadios posteriores del desarrollo de la plántula se observó un incremento en el nivel de transcritos de asparraginasa en hipocótilo en paralelo con una disminución de los niveles de transcrito de asparragina sintetasa. Este cambio en el patrón de expresión de ambos genes coincidió con la fase del desarrollo en la que las reservas del megagametofito se han consumido completamente y se produce una transformación del hipocótilo desde un órgano poco vascularizado a una estructura con un elevado desarrollo de un sistema vascular altamente lignificado. Lo que sugiere que el N almacenado en forma de asparragina, movilizado por la ASPG, juega un papel importante en la fase de transición tras el agotamiento de las reservas del megagametofito (Cañas *et al.*, 2007). De las tres ASPG de pino identificadas en este trabajo, *PpASPG1* es la enzima más

apropiada para cumplir este papel, debido a su mayor afinidad para asparragina como sustrato. No obstante, se ha sugerido una posible redundancia en la función entre las ASPG de los dos subgrupos en *A. thaliana* ya que sus patrones de expresión se solapan (Bruneau *et al.*, 2006). Como las tres ASPG de pino tienen la capacidad de hidrólisis de asparragina, analizamos sus niveles de transcritos mediante qRT-PCR con el objetivo de obtener indicios sobre su implicación en el metabolismo global del nitrógeno de la plántula durante su desarrollo.

Los valores de amplificación de los tres genes sugieren diferencias importantes en los niveles de transcritos (Figura 4.22). El gen *PpASPG3* mostró una baja abundancia de mRNA en todos los órganos y estadios de desarrollo, probablemente debido a que su papel reside en una función general de hidrólisis de dipéptidos de isoaspartilo en las fases finales de la degradación de proteínas, tal y como sugiere también su preferencia por β -Asp-Ala como sustrato. Los niveles relativos de mensajeros de *PpASPG1* y *PpASPG2* son similares en los cotiledones de todos los estadios analizados, así como en el hipocótilo en los dos primeros estadios. Sin embargo, de forma similar a como se describió previamente para *P. sylvestris* (Cañas *et al.*, 2006), los altos niveles de mensajero del gen *PpASPG1* en hipocótilos de los estadios E5 y E6 y en todos los estadios del desarrollo de radículas, junto con las características cinéticas de la enzima que codifica, descritas en este trabajo, sugieren que dicho gen juega un papel principal en la movilización del nitrógeno y el carbono de la asparragina en estos órganos, los cuales son sumideros importantes de nitrógeno y carbono para mantener su desarrollo. La raíz es un órgano en rápida y constante elongación inmediatamente después de la germinación, por lo que requiere un continuo aporte de esqueletos carbonados y nitrógeno para el mantenimiento de su crecimiento. En *Pinus banksiana*, aspartato y asparragina juntos constituyen los aminoácidos libres de mayor abundancia en brotes y también en raíces (Cyr *et al.*, 1990), lo cual apoya un papel cuantitativamente relevante de la ASPG en los flujos de nitrógeno a los órganos sumideros en pino. En *P. sylvestris* se detectaron altos niveles de expresión del gen ortólogo de *PpASPG1* en radícula 8 días después de la

imbibición (Cañas *et al.*, 2007), lo que sugiere que la asparragina puede ser uno de los vehículos utilizados para proveer a la raíz de ambos recursos. Por otro lado, la abundancia de mRNA en hipocótilo aumenta en los estadios coincidentes con el agotamiento de las reservas de la semilla. Por lo que la plántula acude a las reservas en forma de asparragina acumuladas en este órgano a partir del nitrógeno no utilizado de las reservas durante los primeros estadios postgerminativos. Entre 12 y 30 días después de la imbibición de semillas de *P. sylvestris* se observó el cese de la elongación del hipocótilo y la aparición de cambios estructurales tales como un gran desarrollo del sistema vascular, incluyendo la unificación de los haces de elementos de xilema en un anillo continuo (Cañas *et al.*, 2007). Durante esta transición se observó un aumento en la abundancia de transcritos para la ASPG, sugiriendo que el hipocótilo consume las reservas de asparragina para proveer N y C para asegurar material suficiente para el desarrollo de un sistema vascular que será esencial para el transporte eficiente de agua y nutrientes minerales una vez agotadas las reservas de la semilla. Por lo que resulta un proceso esencial para la adaptación y supervivencia de las plántulas tras la germinación. Plantas mutantes para el gen *NSE1* de *Lotus japonicus*, deficientes en asparraginasa dependiente de potasio, tenían una longitud de tallo reducida (Credali *et al.*, 2013), sugiriendo la importancia de esta enzima en el mantenimiento de la tasa de crecimiento del tallo en esta especie.

La menor abundancia de transcritos *PpASPG1* en el ápice de la raíz (Figura 4.23), en el que se localiza el meristemo radicular, contrasta con el perfil de expresión en especies angiospermas del género *Lupinus* dónde se detecta la ASPG dependiente de potasio solamente en el primer centímetro de la raíz (Chang and Farnden, 1981). Los diferentes tipos celulares se diferencian mientras aumenta la distancia de la célula con respecto a este meristemo (Lucas *et al.*, 2013). El xilema primario se forma a partir de células provasculares cerca del meristemo apical y el xilema secundario se diferencia a partir de las células de la zona cambial (Côté *et al.*, 2010). En gimnospermas y dicotiledóneas durante el crecimiento secundario, cuando ya está establecida la vasculatura primaria, el procambium da lugar al cambium que es un

meristemo lateral que induce el incremento en biomasa en dirección radial generando el xilema secundario (Ruzicka *et al.*, 2015). Transcritos de la ASPG de *Pinus sylvestris* se detectaron particularmente en la región del cambium vascular de plántulas jóvenes (Cañas *et al.*, 2007). La abundancia de mRNA de *PpASPG1* a lo largo de la radícula con la excepción del ápice (Figura 4.23) apoya una función para este gen asociada al sistema vascular.

Por otro lado, los niveles relativos de mRNA de *PpASPG1* y *PpASPG2* a lo largo de los tallos de plántulas de dos años (Figura 4.24) muestran que ambos genes siguen activos en tallos en estadios más avanzado de desarrollo. En particular, la presencia de transcritos de *PpASPG1* en árboles adultos (Figuras 4.25 y 4.26) muestra la funcionalidad de este gen a lo largo de todo el desarrollo del árbol. Cuando comparamos muestras de 4 individuos diferentes (con más de 20 años de edad) muestreados en marzo y mayo observamos la ausencia de un patrón consistente en expresión entre los diferentes individuos. Estos resultados sugieren una alta plasticidad en la regulación de la transcripción de este gen en coníferas en respuesta a factores internos y/o externos.

5.4.2 Implicación de la asparraginasa en la adaptación del metabolismo de pino a cambios ambientales

La variabilidad en el patrón de expresión de *PpASPG1* en individuos adultos sugiere una plasticidad en la expresión de este gen que podría estar relacionado con un papel en la adaptación metabólica en respuesta a cambios internos y externos. Para contrastar esta hipótesis analizamos la respuesta de los genes de asparraginasa de *P. pinaster* a distintos efectos externos.

Los bajos niveles de transcrito de *PpASPG1* en los distintos órganos de plántulas en ausencia de nitrógeno en la solución de riego y el significativo aumento tras 24 horas de suministro de $(\text{NH}_4)_2\text{SO}_4$ (Figura 4.27), muestran la importancia de la asparragina y la asparraginasa en la readaptación inicial de

la planta a cambios en la disponibilidad de amonio. Por otro lado, la ausencia de cambios relevantes en los niveles de transcrito del gen de la isoforma 1b de la glutamina sintetasa (GS1b) en las distintas condiciones de disponibilidad de nitrógeno ensayadas, a excepción de en cotiledones expuestos a una prolongada deficiencia de nitrógeno, así como la respuesta de los niveles de mensajeros del gen de ornitina- δ -aminotransferasa (δ -OAT) similar a la observada para los mensajeros de *PpASPG1* sugiere que *PpASPG1* y el gen que codifica δ -OAT parecen ser claves para la adaptación inicial de la plántulas a cambios en las condiciones de disponibilidad de nitrógeno. Cañas *et al.*, (2008) ya había sugerido un papel para el gen δ -OAT en el desarrollo de plántulas de pino, y además, en *A. thaliana* se demostró que la función del gen δ -OAT está asociada al catabolismo de arginina (Funck *et al.*, 2008), otro aminoácido que puede actuar como reserva temporal de nitrógeno.

Por el contrario, los perfiles de expresión de los genes *PpASGP2* y *PpASPG3* en respuesta a la variación en la disponibilidad de amonio mostraron una tendencia similar entre ellos y opuesta a la *PpASPG1* (Figura 4.27 y 4.28). En general, los niveles de transcritos de ambos genes disminuyeron bajo la condición de disponibilidad moderada y alta de amonio, siendo más altos bajo deficiencia de amonio. Estos perfiles de expresión apoyan una función diferente para ambos genes respecto a *PpASPG1*. En particular, la alta disponibilidad de amonio puede disminuir la actividad de recuperación de aminoácidos aberrantes en forma de dipéptidos de aspartilo, mientras que, en ausencia de un suministro de nitrógeno, maximizar la recuperación y reciclado de todo el nitrógeno endógeno es vital para la supervivencia de la plántula. La observación de que tras 29 días los niveles de ambos genes son inferiores en plántulas regadas con una solución de 10 mM de $(\text{NH}_4)_2\text{SO}_4$ que en aquellas regadas con 3 mM de $(\text{NH}_4)_2\text{SO}_4$ apoyan esta posibilidad.

Una rápida adaptación al estrés requiere una determinada plasticidad metabólica que permita al organismo proporcionar los componentes de respuesta requeridos y la ASPG podría tener una función en movilizar nitrógeno para sostener las vías metabólicas adaptativas. Análisis

transcriptómicos en *A. thaliana* muestran un cambio significativo en el nivel de expresión del gen de la ASPG dependiente de potasio en respuesta a factores de estrés y la respuesta es dependiente de tejido (Aceituno *et al.*, 2008). En hojas de soja se detectó un aumento en el nivel de mRNA de ASPG después de un tratamiento con ácido abscísico, frío y sal (Cho *et al.*, 2007) y en algodón el gen se indujo por un ataque con patógenos (Dubey *et al.*, 2013). En especies gimnospermas la asparraginasa puede estar también relacionada con factores de estrés, ya que en células de *Taxus chinensis* tratadas con metil jasmonato se observó un aumento en la expresión de un gen de asparraginasa, junto con factores de transcripción relacionados con la respuesta a estrés (Li *et al.*, 2012). Puesto que las plántulas jóvenes son muy vulnerables a cambios en las condiciones ambientales, se evaluó si en ellas los genes de ASPG respondían en respuesta a condiciones adversas.

Se detectaron respuestas significativas en el nivel de expresión de *PpASPG1* frente a todos los tratamientos de estrés ensayados, de una manera dependiente del órgano y del tiempo (Figura 4.30 a 4.33). El mayor aumento de transcritos se registró en respuesta al jasmonato, una hormona que aparte de inducir una reprogramación en respuesta a patógenos, también está implicada en varios procesos de desarrollo entre los cuales está el desarrollo de la raíz (Pauwels *et al.*, 2008). En tallos de árboles jóvenes de *Picea abies* esta hormona cambia el programa de desarrollo del cambium e induce la diferenciación *de novo* de conductos de resina en xilema (Martin *et al.*, 2002). Hudgins and Franceschi (2004) concluyeron que esta reprogramación del cambium era resultado de la acción del etileno, producto de la actividad ACC oxidasa que es estimulada a su vez por jasmonato y por heridas. Por ende, la respuesta del gen *PpASPG1* de pino al jasmonato podría ser indirecta e incorpora más variables, tales como amplificación de la señal y retraso en el tiempo en la respuesta entre los diferentes órganos. Diferentes poblaciones de pinos controlan su metabolismo de jasmonatos de manera diferente, e incluso entre individuos y entre tejidos dentro de un mismo individuo se encuentran composiciones moleculares diferentes de jasmonatos, por lo cual esta familia de moléculas constituye un complejo de señalización muy flexible (Pedranzani

et al., 2007). El patrón de expresión de *PpASPG1* observado en hipocótilos en respuesta a heridas es comparable con el de jasmonato aunque el efecto de inducción es menor. Las defensas de coníferas frente a patógenos incluye, aparte de cambios anatómicos, la síntesis de resinas, compuestos fenólicos y alcaloides y el efecto de cada tipo de respuesta depende de la dosis de jasmonato (Moreira *et al.*, 2012). En *A. thaliana* los jasmonatos inducen la síntesis de alcohol coniferílico y estimulan genes de síntesis de aminoácidos aromáticos, del metabolismo de amonio y las rutas biosintéticas de serina y cisteína (Pauwels *et al.*, 2008). Considerando la participación del aspartato en varias rutas metabólicas relacionadas con los procesos de adaptación, observados tanto en especies gimnospermas como angiospermas, lógicamente la ASPG debe tener un papel importante en su suministro. Por ejemplo, SAM sintetasa aumenta el nivel de expresión en respuesta a varios factores de estrés al igual que la treonina desaminasa (TD) que se induce por metil-jasmonato y heridas en tomate y por patógenos en tabaco (Azevedo *et al.*, 2006). De la TD existen dos isoformas en plantas de las cuales una se relaciona con tejidos jóvenes y tiene una función biosintética mientras que la otra, de función catabólica, se expresa durante la senescencia. La isoforma anabólica es inducida en respuesta a estrés paralelamente a una activación de rutas metabólicas de síntesis de compuestos de defensa. Lógicamente el consumo tanto de treonina como de metionina requiere de un aporte considerable de aspartato. Por otro lado, el catabolismo de los aminoácidos de la familia del aspartato contribuye, a través de un nexo regulatorio con el ciclo del ácido cítrico, a proveer energía para responder a factores de estrés que inducen su déficit (Galili, 2011). El hecho de que las heridas provocadas en cotiledones inducen una bajada en actividad L-asparaginasa en radículas podría parar la movilización de N en este órgano menos propenso a ser atacado y preservar las reservas de C y N en forma de asparagina o incluso transportarlas hacia los órganos que precisan de C y N para reforzar sus sistemas de defensa.

El tratamiento con ABA sin embargo produjo un efecto contrario al jasmonato en el nivel de mensajeros del gen *PpASPG1* (Figura 4.30 y 4.32). En hojas de

A. thaliana la ASPG dependiente de potasio se expresa altamente en células guarda de los estomas y la cantidad de transcritos también se reduce por ABA 50 μM a 3 horas después del tratamiento en estas células (Yang *et al.*, 2008). Los pinos, a diferencia de algunas otras coníferas, acumulan altos niveles de ABA durante periodos de sequía por lo cual se le asocia a la hormona en pino también una función relacionada con el cierre de los estomas (Brodribb *et al.*, 2014). Se ha descrito que el ABA induce cambios fisiológicos y morfológicos que mejoran la tolerancia a la sequía en *Pinus banksiana* y reducen la elongación del brote (Marshall *et al.*, 1991). En *Picea abies* también se observó una elongación reducida del brote por efecto de esta hormona (Heide, 1986). Tanto jasmonato como ABA están relacionadas, entre otros procesos, con respuestas a factores de estrés. Sin embargo, tienen un efecto contrario sobre los transcritos de *PpASPG1*, sugiriendo que la respuesta del gen está asociada a las necesidades de la planta para poder responder a las características específicas del estrés. ABA se acumula en condiciones de estrés por sequía y estrés salino (Lee *et al.*, 2005) y en nuestros experimentos tanto el tratamiento con ABA como estrés salino (Figura 4.32 y 4.33) producen una disminución en los niveles de transcritos del gen *PpASPG1*. Posiblemente, la tasa de síntesis de proteínas en los tejidos sumideros se ve reducida debido a la baja disponibilidad de esqueletos de carbono y la reducida actividad metabólica como resultado de una disminución en la tasa de fotosíntesis asociada al cierre de los estomas.

El gen *PpASPG2* también responde a los tratamientos con las hormonas jasmonato y ABA además de heridas, generalmente con un aumento en la cantidad de transcritos (Figura 4.30 a 4.33), por lo que podría jugar un papel también en la respuesta al estrés. Sin embargo, los bajos valores de V_{max} para ambos sustratos ensayados con la proteína recombinante (L-Asn y β -Asp-Ala) y la falta de preferencia por uno de los dos sustratos no parece indicar que juega un papel esencial en la respuesta de adaptación del metabolismo del nitrógeno a la situación de estrés. No obstante, podría ser que tuviera una alta afinidad por un dipéptido de isoaspartilo no ensayado y que en condiciones de estrés se incrementara la expresión del gen y aumentase la abundancia de la

proteína, como parte de un proceso general de degradación y reciclaje de proteínas para disponer de material suficiente para la síntesis de proteínas requeridas en la respuesta al estrés, en condiciones donde todos los recursos deben ser maximizados para la supervivencia.

El perfil de expresión de *PpASPG1* apoya un papel del mismo en la adaptación metabólica de la planta de pino como parte de las respuestas iniciales a condiciones cambiantes, por lo que sería relevante para una especie con ciclos de vida tan extensos.

6. CONCLUSIONES

1. El transcriptoma del pino marítimo (*Pinus pinaster*) contiene un total de tres mRNA diferentes cuyas secuencias codifican para asparraginasas. Considerando las estructuras primarias de las proteínas, PpASPG1 se clasifica en el grupo dependiente de potasio mientras que PpASPG2 y PpASPG3 son más parecidas al tipo independiente de potasio.
2. Las tres asparraginasas de pino presentan una doble actividad asparraginasa/isoaspartilo dipeptidasa pero difieren en su especificidad de sustrato: PpASPG1 presenta una preferencia por L-asparragina como sustrato, mientras que PpASPG3 utiliza preferentemente β -Asp-Ala y PpASPG2 no mostró una preferencia particular por ninguno de los dos sustratos.
3. De los tres genes de asparraginasa de pino, las características enzimáticas y perfil de expresión del gen apuntan a *PpASPG1* como la forma principalmente implicada en la provisión de amonio y aspartato a partir de asparragina para responder a las necesidades metabólicas requeridas por procesos de desarrollo o en respuesta a factores externos.
4. La región de 77 aminoácidos del extremo carboxilo de la subunidad α , denominada región variable por su alta variabilidad en longitud y secuencia de aminoácidos entre asparraginasas, parece ser determinante de la preferencia de sustrato mostrada por PpASPG1, además de contribuir a la estabilidad de la enzima frente a la temperatura.
5. En el caso de PpASPG1 esta región variable tiene además un efecto sobre la maduración de la enzima: influye positivamente en la eficiencia del proceso de maduración de la proteína *in planta* pero impide la capacidad autocatalítica del precursor *in vitro*, por lo que podría formar parte de un mecanismo de regulación postraducciona en el que estarían implicados otros factores desconocidos.

6. El ion potasio tiene un efecto termoprotector sobre PpASPG1 pero su efecto activador es más limitado de lo que se ha descrito para las asparraginasas del tipo dependiente de potasio en especies angiospermas, pudiendo el aumento en V_{\max} ser resultado de una mayor estabilidad de la enzima.
7. La región promotora proximal del gen *PpASPG1* de *P. pinaster* presenta un alto grado de conservación con la del gen ortólogo de *P. sylvestris* y contiene elementos identificados *in silico* como cajas *cis* relacionadas con regulación de la transcripción en respuesta a varios factores de estrés, los cuales se han confirmado experimentalmente que afectan a los niveles de transcritos del gen. Por el contrario, la región distal del promotor presenta una mayor variabilidad, pudiendo incluir elementos que son también importantes para modular la expresión del gen.
8. La región promotora proximal presenta una variabilidad intraespecífica media de 1 SNP por cada 83 pares de bases en 11 procedencias diferentes de *P. pinaster*, afectando a posibles elementos *cis* de regulación transcripcional y pudiendo contribuir, junto con las variaciones observadas en la región distal, a diferencias entre procedencias en las respuestas a diversos factores externos.

7. BIBLIOGRAFÍA

Aceituno FF, Moseyko N, Rhee SY, Gutiérrez R a. 2008. The rules of gene expression in plants: organ identity and gene body methylation are key factors for regulation of gene expression in *Arabidopsis thaliana*. *BMC genomics* **9**, 438.

Acosta-Martinez V, Moore-Kucera J, Cotton J, Gardner T, Wester D. 2014. Soil enzyme activities during the 2011 Texas record drought/heat wave and implications to biogeochemical cycling and organic matter dynamics. *Applied Soil Ecology* **75**, 43–51.

Ali EMM. 2009. Purification and characterization of *Vigna unguiculata* cultivar asparaginase. *The egyptian journal of biochemistry and molecular biology* **2**, 145–161.

Alía R, García del Barrio JM, Iglesias Sauce S, Mancha Núñez JA, de Miguel y del Ángel J, Nicolás Peragón JL, Pérez Martín F, Sánchez de Ron D. 2009. *Regiones de procedencia de especies forestales en España*.

Atkins CA, Pate JS, Sharkey PJ. 1975. Asparagine metabolism - key to the nitrogen nutrition of developing legume seeds. *Plant Physiol.* **56**, 807–812.

Avila C, Suárez MF, Gómez-Maldonado J, Cánovas FM. 2001. Spatial and temporal expression of two cytosolic glutamine synthetase genes in Scots pine: functional implications on nitrogen metabolism during early stages of conifer development. *Plant Journal* **25**, 93–102.

Azevedo RA, Lancien M, Lea PJ. 2006. The aspartic acid metabolic pathway, an exciting and essential pathway in plants. *Amino Acids* **30**, 143–162.

Bejger M, Imiolczyk B, Clavel D, Gilski M, Pajak A, Marsolais F, Jaskolski M. 2014. Na⁺/K⁺ exchange switches the catalytic apparatus of potassium-dependent plant L-asparaginase. *Acta Crystallographica Section D: Biological Crystallography* **70**, 1854–1872.

Biasini M, Bienert S, Waterhouse A, et al. 2014. SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research* **42**, 252–258.

Borek D, Jaskolski M. 2000. Crystallization and preliminary crystallographic studies of a new L-asparaginase encoded by the *Escherichia coli* genome. *Acta Crystallographica Section D*. **56**, 1505–1507.

Borek D, Jaskólski M. 2001. Sequence analysis of enzymes with asparaginase activity. *Acta Biochimica Polonica* **48**, 893–902.

Borek D, Michalska K, Brzezinski K, Kisiel A, Podkowinski J, Bonthron DT, Krowarsch D, Otlewski J, Jaskólski M. 2004. Expression, purification and catalytic activity of *Lupinus luteus* asparagine β -amidohydrolase and its *Escherichia coli* homolog. *European Journal of Biochemistry* **271**, 3215–3226.

Bradford M. 1976. A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254.

Brannigan J, Dodson G, Duggleby H, Moody P, Smith J, Tomchick D, Murzin A. 1995. A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature* **378**, 416–419.

Brodrribb TJ, McAdam SAM, Jordan GJ, Martins SC V. 2014. Conifer species adapt to low-rainfall climates by following one of two divergent pathways. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 14489–14493.

Broome JD. 1963. Evidence that the L-asparaginase of guinea pig serum is responsible for its antilymphoma effects. *J. Exp. Med* **Jul**, 99–120.

Bruneau L, Chapman R, Marsolais F. 2006. Co-occurrence of both L-asparaginase subtypes in *Arabidopsis*: At3g16150 encodes a K⁺-dependent L-asparaginase. *Planta* **224**, 668–679.

Canales J, Ávila C, Cantón FR, Pacheco-Villalobos D, Díaz-Moreno S, Ariza D, Molina-Rueda JJ, Navarro-Cerrillo RM, Claros MG, Cánovas FM. 2012. Gene expression profiling in the stem of young maritime pine trees: detection of ammonium stress-responsive genes in the apex. *Trees - Structure and Function* **26**, 609–619.

Canales J, Bautista R, Label P, et al. 2014. De novo assembly of maritime pine transcriptome: implications for forest breeding and biotechnology. *Plant Biotechnology Journal* **12**, 286–299.

Canales J, Flores-Monterrosso A, Rueda-López M, Avila C, Cánovas FM. 2010. Identification of genes regulated by ammonium availability in the roots of maritime pine trees. *Amino Acids* **39**, 991–1001.

Cánovas FM, Avila C, Cantón FR, Cañas RA, De La Torre F. 2007. Ammonium assimilation and amino acid metabolism in conifers. *Journal of Experimental Botany* **58**, 2307–2318.

Cánovas FM, Cantón FR, García-Gutiérrez A, Gallardo F, Crespillo R. 1998. Molecular physiology of glutamine and glutamate biosynthesis in developing seedlings of conifers. *Physiologia Plantarum* **103**, 287–294.

Cantón FR, Suárez MF, Cánovas FM. 2005. Molecular aspects of nitrogen mobilization and recycling in trees. *Photosynthesis Research* **83**, 265–278.

Cantor JR, Stone EM, Chantranupong L, Georgiou G. 2009. The human asparaginase-like protein 1 hASRGL1 is an Ntn hydrolase with β -aspartyl peptidase activity. *Biochemistry* **48**, 11026–11031.

Cañas RA, De La Torre F, Cánovas FM, Cantón FR. 2006. High levels of asparagine synthetase in hypocotyls of pine seedlings suggest a role of the enzyme in re-allocation of seed-stored nitrogen. *Planta* **224**, 83–95.

Cañas RA, De La Torre F, Cánovas FM, Cantón FR. 2007. Coordination of *PsAS1* and *PsASPG* expression controls timing of re-allocated N utilization in hypocotyls of pine seedlings. *Planta* **225**, 1205–1219.

Cañas RA, Villalobos DP, Díaz-Moreno SM, Cánovas FM, Cantón FR. 2008. Molecular and functional analyses support a role of Ornithine- δ -aminotransferase in the provision of glutamate for glutamine biosynthesis during pine germination. *Plant physiology* **148**, 77–88.

Carter P, Wells JA. 1988. Dissecting the catalytic triad of a serine protease. *Nature* **332**, 564–568.

Di Cera E. 2006. A structural perspective on enzymes activated by monovalent cations. *Journal of Biological Chemistry* **281**, 1305–1308.

Chagas EP, Sodek L. 2001. Purification and properties of asparaginase from the testa of immature seeds of pea (*Pisum sativum* L.). *Brazilian Archives of Biology and Technology* **44**, 239–246.

Chang KS, Farnden KJF. 1981. Purification and properties of asparaginase from *Lupinus arboreus* and *Lupinus angustifolius*. *Archives of Biochemistry and Biophysics* **208**, 49–58.

Chen C, Winans SC. 1991. Controlled expression of the transcriptional activator gene *virG* in *Agrobacterium tumefaciens* by using the *Escherichia coli* lac promoter. *Journal of bacteriology* **173**, 1139–1144.

Cho CW, Lee HJ, Chung E, et al. 2007. Molecular characterization of the soybean L-asparaginase gene induced by low temperature stress. *Mol Cells* **23**, 280–286.

Clementi A. 1922. La desamidation enzymatique de l'asparagine chez les differentes especes animals et la signification physiologique de sa presence dans l'organisme. *Arch. Intern. Physiol.* **19**, 369–398.

Coleman GD, Banados MP, Chen T. 1994. Poplar bark storage protein and a related wound-induced gene are differentially induced by nitrogen. *Plant physiology* **106**, 211–215.

Coruzzi GM. 2003. Primary N-assimilation into amino acids in *Arabidopsis*. *The Arabidopsis Book* **2**, 1–17.

Côté CL, Boileau F, Roy V, Ouellet M, Levasseur C, Morency M-J, Cooke JEK, Séguin A, MacKay JJ. 2010. Gene family structure, expression and functional analysis of HD-Zip III genes in angiosperm and gymnosperm forest trees. *BMC plant biology* **10**, 273.

Credali A, Díaz-Quintana A, García-Calderón M, De la Rosa MA, Márquez AJ, Vega JM. 2011. Structural analysis of K⁺ dependence in L-asparaginases from *Lotus japonicus*. *Planta* **234**, 109–122.

Credali A, García-Calderón M, Dam S, Perry J, Díaz-Quintana A, Parniske M, Wang TL, Stougaard J, Vega JM, Márquez AJ. 2013. The K⁺-dependent asparaginase, NSE1, is crucial for plant growth and seed production in *Lotus japonicus*. *Plant and Cell Physiology* **54**, 107–118.

Cyr DR, Buxton GF, Webb DP, Dumbroff EB. 1990. Accumulation of free amino acids in the shoots and roots of three northern conifers during drought. *Tree Physiology* **6**, 293–303.

Dellaporta SL, Wood J, Hicks JB. 1983. A plant DNA miniprep: Version II. *Plant Molecular Biology Reporter* **1**, 19–21.

Díaz-Moreno SM. 2010. Identificación de genes determinantes de caracteres de la madera juvenil y madura mediante análisis transcriptómicos en pino. Universidad de Málaga, Spain.

Dickson JMJJ, Vincze E, Grant MR, Smith LA, Rodber KA, Farnden KJF, Reynolds PHS. 1992. Molecular cloning of the gene encoding developing seed L-asparaginase from *Lupinus angustifolius*. *Plant Molecular Biology* **20**, 333–336.

Dubey NK, Goel R, Ranjan A, Idris A, Singh SK, Bag SK, Chandrashekar K, Pandey KD, Singh PK, Sawant S V. 2013. Comparative transcriptome analysis of *Gossypium hirsutum* L. in response to sap sucking insects: aphid and whitefly. *BMC genomics* **14**, 241.

Duggleby H, Tolley S, Hill C, Dodson E, Dodson G, Moody P. 1995. Penicillin acylase has a single-amino-acid catalytic centre. *Nature* **373**, 264–268.

Epp O, Steigemann W, Formanek H, Huber R. 1971. Crystallographic evidence for the tetrameric subunit structure of L-asparaginase from *Escherichia coli*. *European journal of biochemistry / FEBS* **20**, 432–437.

Flannigan MD, Logan KA, Amiro BD, Skinner WR, Stocks BJ. 2005. Future area burned in Canada. *Climatic Change* **72**, 1–16.

Flores-Monterroso A, Canales J, de la Torre F, Ávila C, Cánovas FM. 2013. Identification of genes differentially expressed in ectomycorrhizal roots during the *Pinus pinaster*-*Laccaria bicolor* interaction. *Planta* **237**, 1637–1650.

Frankenberger WT, Tabatabai MA. 1991. L-Asparaginase activity of soils. *Biology and Fertility of Soils* **11**, 6–12.

Friesen N, Brandes A, Heslop-Harrison JS. 2001. Diversity, origin, and distribution of retrotransposons (gypsy and copia) in conifers. *Molecular biology and evolution* **18**, 1176–1188.

Funck D, Stadelhofer B, Koch W. 2008. Ornithine- δ -aminotransferase is essential for arginine catabolism but not for proline biosynthesis. *BMC plant biology* **8**.

Gabriel M, Telmer PG, Marsolais F. 2012. Role of asparaginase variable loop at the carboxyl terminal of the alpha subunit in the determination of substrate preference in plants. *Planta* **235**, 1013–1022.

Galiano L, Martínez-Vilalta J, Lloret F. 2010. Drought-induced multifactor decline of Scots pine in the Pyrenees and potential vegetation change by the expansion of co-occurring oak species. *Ecosystems* **13**, 978–991.

Galili G. 2011. The aspartate-family pathway of plants. Linking production of essential amino acids with energy and stress regulation. *Plant signaling & behavior* **6**, 192–195.

Galperin MY, Koonin E V. 2012. Divergence and convergence in enzyme evolution. *Journal of Biological Chemistry* **287**, 21–28.

Gaufichon L, Masclaux-Daubresse C, Tcherkez G, et al. 2013. *Arabidopsis thaliana* *ASN2* encoding asparagine synthetase is involved in the control of nitrogen assimilation and export during vegetative growth. *Plant, Cell and Environment* **36**, 328–342.

Grant M, Bevan MW. 1994. Asparaginase gene expression is regulated in a complex spatial and temporal pattern in nitrogen-sink tissues. *The Plant Journal* **5**, 695–704.

Gutiérrez RA, Lejay L V, Dean A, Chiaromonte F, Shasha DE, Coruzzi GM. 2007. Qualitative network models and genome-wide expression data define carbon/nitrogen-responsive molecular machines in *Arabidopsis*. *Genome biology* **8**, R7.

Hanson AD, Roje S. 2001. One-carbon metabolism in higher plants. *Annu. Rev. Plant Physiol.* **52**, 119–137.

Heide OM. 1986. Effects of ABA application on cessation of shoot elongation in long-day grown Norway spruce seedlings. *Tree physiology* **1**, 79–83.

Hejazi M, Piotukh K, Mattow J, Deutzmann R, Volkmer-Engert R, Lockau W. 2002. Isoaspartyl dipeptidase activity of plant-type asparaginases. *The Biochemical journal* **364**, 129–136.

Higo K, Ugawa Y, Iwamoto M, Korenaga T. 1999. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res* **27**, 297–300.

Hildebrandt TM, Nesi AN, Araújo WL, Braun H. 2015. Amino acid catabolism in plants. *MOLECULAR PLANT* **8**, 1563–1579.

Ho DHW, Whitecar JP, Luce JK, Frei E. 1970. L-asparagine requirement and the effect of L-asparaginase on the normal and leukemic human bone marrow. *Cancer Research* **30**, 466–473.

Hudgins JW, Franceschi VR. 2004. Methyl jasmonate-induced ethylene production is responsible for conifer phloem defense responses and reprogramming of stem cambial zone for traumatic resin duct formation. *Plant physiology* **135**, 2134–2149.

Ireland RJ, Joy KW. 1983. Subcellular localization of asparaginase and asparagine aminotransferase in *Pisum sativum* leaves. *Plant Physiol.* **72**, 1127–1129.

Islam-faridi N, Peterson DG, Majid MA, Magbanua ZV, Nelson CD. 2009. *Chromosomal Distribution of Repetitive DNA Sequences in Pine using Fluorescent In Situ Hybridization (FISH)*. Southern Research Station (U.S. Forest Service)

Ivanov A, Kameka A, Pajak A, Bruneau L, Beyaert R, Hernández-Sebastià C, Marsolais F. 2012. *Arabidopsis* mutants lacking asparaginases develop normally but exhibit enhanced root inhibition by exogenous asparagine. *Amino Acids* **42**, 2307–2318.

Kader MA, Lindberg S. 2010. Cytosolic calcium and pH signaling in plants under salinity stress. *Plant signaling & behavior* **5**, 233–238.

Kamm A, Doudrick RL, Heslop-Harrison JS, Schmidt T. 1996. The genomic and physical organization of *Ty1-copia*-like sequences as a component of large genomes in *Pinus elliottii* var. *elliottii* and other gymnosperms. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 2708–2713.

Khalaf ZA. 2012. Extraction and Purification of Asparaginase enzyme from *Pisum sativum* plant and studying their cytotoxicity against L20B tumor cell line. Al-Nahrain University, Republic of Iraq.

Kidd JG. 1953. Regression of transplanted lymphomas induced in vivo by means of normal guinea pig serum. *J. Exp Med* **Dec**, 565–582.

King JE, Gifford DJ. 1997. Amino acid utilization in seeds of Loblolly pine during germination and early seedling growth. *Plant Physiol.*, 1125–1135.

Kirma M, Araújo WL, Fernie AR, Galili G. 2012. The multifaceted role of aspartate-family amino acids in plant metabolism. *Journal of Experimental Botany* **63**, 4995–5001.

Kronzucker HJ, Siddiqi MY, Glass ADM. 1997. Conifer root discrimination against soil nitrate and the ecology of forest succession. *Nature* **385**, 59–61.

Kurz WA, Dymond CC, Stinson G, Rampley GJ, Neilson ET, Carroll AL, Ebata T, Safranyik L. 2008. Mountain pine beetle and forest carbon feedback to climate change. *Nature* **452**, 987–990.

de la Torre F, El-Azaz J, Avila C, Cánovas FM. 2014. Deciphering the role of aspartate and prephenate aminotransferase activities in plastid nitrogen metabolism. *Plant physiology* **164**, 92–104.

Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

Lam H-M, Coschigano K, Schultz C, Melo-Oliveira R, Tjaden G, Oliveira I, Ngai N, Hsieh M-H, Coruzzi G. 1995. Use of Arabidopsis mutants and genes to study amide amino acid biosynthesis. *The Plant Cell* **7**, 887.

Lancien M, Martin M, Hsieh M, Leustek T, Goodman H, Coruzzi GM. 2002. Arabidopsis *glt1* -T mutant defines a role for NADH-GOGAT in the non-photorespiratory ammonium assimilatory pathway. *The Plant Journal* **29**, 347–358.

Lang S. 1904. Über desamidierung im Tierkörper. *Beitr. Chem. Physiol. Pathol.* **5**, 321–345.

Lea PJ, Festenstein GN, Hughes JS, Mifflin BJ. 1984. An immunological and enzymological survey of asparaginase in seeds of *Lupinus*. *Phytochemistry* **23**, 511–514.

Lea PJ, Mifflin BJ. 2003. Glutamate synthase and the synthesis of glutamate in plants. *Plant physiology and Biochemistry* **41**, 555–564.

Lea PJ, Sodek L, Parry MAJ, Shewry PR, Halford NG. 2007. Asparagine in plants. *Annals of Applied Biology* **150**, 1–26.

Lee BH, Henderson DA, Zhu JK. 2005. The *Arabidopsis* cold-responsive transcriptome and its regulation by ICE1. *Plant Cell* **17**, 3155–3175.

Lees EM, Blakeney AB. 1970. The distribution of asparaginase activity in legumes. *BBA - General Subjects* **215**, 145–151.

Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S. 2002. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic acids research* **30**, 325–327.

Lev-yadun S, Sederoff R. 2000. Pines as model gymnosperms to study evolution , wood formation , and perennial growth. *J Plant Growth Regul* **19**, 290–305.

Li S, Zhang P, Zhang M, Fu C, Zhao C, Dong Y, Guo A, Yu L. 2012. Transcriptional profile of *Taxus chinensis* cells in response to methyl jasmonate. *BMC Genomics* **13**, 295.

Liao Z, Chen M, Guo L, Gong Y, Tang F, Sun X, Tang K. 2004. Rapid isolation of high-quality total RNA from taxus and ginkgo. *Preparative biochemistry & biotechnology* **34**, 209–214.

Lindbo JA. 2007. TRBO: A high-efficiency *Tobacco mosaic virus* RNA-based overexpression vector. *Plant Physiol.* **145**, 1232–1240.

Liu J-J, Sniezko RA, Sturrock RN, Chen H. 2014. Western white pine SNP discovery and high-throughput genotyping for breeding and conservation applications. *BMC plant biology* **14**, 380.

Loan DK, Con TH, Hong TT, Mai Ly LT. 2013. Quick determination of ammonia ions in water environment based on thymol color creating reaction. *Environmental Sciences* **1**, 83–92.

Lough TJ, Chang KS, Carne A, Monk BC, Reynolds PHS, Farnden KJF. 1992. L-asparaginase from developing seeds of *Lupinus arboreus*. *Phytochemistry* **31**, 1519–1527.

Lucas WJ, Groover A, Lichtenberger R, et al. 2013. The plant vascular system: evolution, development and functions. *Journal of Integrative Plant Biology* **55**, 294–388.

Mackay J, Dean JFD, Plomion C, et al. 2012. Towards decoding the conifer giga-genome. *Plant Molecular Biology* **80**, 555–569.

Marshall JG, Scarratt JB, Dumbroff EB. 1991. Induction of drought resistance by abscisic acid and paclobutrazol in jack pine. *Tree Physiology* **8**, 415–421.

Martin D, Tholl D, Gershenzon J. 2002. Methyl jasmonate induces traumatic resin ducts, terpenoid resin biosynthesis, and terpenoid accumulation in developing xylem of Norway spruce stems. *Society* **129**, 1003–1018.

Michalska K, Bujacz G, Jaskólski M. 2006. Crystal structure of plant asparaginase. *Journal of Molecular Biology* **360**, 105–116.

Michalska K, Hernandez-Santoyo A, Jaskolski M. 2008. The mechanism of autocatalytic activation of plant-type L-asparaginases. *Journal of Biological Chemistry* **283**, 13388–13397.

Michalska K, Jaskólski M. 2006. Structural aspects of L-asparaginases, their friends and relations. *Acta Biochimica Polonica* **53**, 627–640.

Ministerio de agricultura alimentación y medio ambiente. 2012. *Avance anuario de estadística forestal 2012*.

Möllerling H. 1985. L-aspartate and L-asparagine. In: Bergmeyer HU,, In: Bergmeyer J,, In: Grass IM, eds. *Methods of Enzymatic Analysis*, Vol. VIII. 3rd edn. VCH Verlagsgesellschaft, Weinheim, 350–357.

Moreira X, Zas R, Sampedro L. 2012. Methyl jasmonate as chemical elicitor of induced responses and anti-herbivory resistance in young conifer trees. *Plant Defence: Biological control*.345–362.

Morse AM, Peterson DG, Islam-faridi MN, et al. 2009. Evolution of genome size and complexity in *Pinus*. *PLoS ONE* **4**, 1–11.

Müntz K. 1998. Deposition of storage proteins. *Plant Molecular Biology* **38**, 77–99.

Murray BG. 1998. Nuclear DNA Amounts in Pteridophytes. *Annals of Botany* **82**, 3–15.

Neale DB, Kremer A. 2011. Forest tree genomics: growing resources and applications. *Nature reviews. Genetics* **12**, 111–122.

Neves LG, Davis JM, Barbazuk WB, Kirst M. 2013. Whole-exome targeted sequencing of the uncharacterized pine genome. *Plant Journal* **75**, 146–156.

Nomme J, Su Y, Konrad M, Lavie A. 2012. Structures of apo and product-bound human L-asparaginase: Insights into the mechanism of autoproteolysis and substrate hydrolysis. *Biochemistry* **51**, 6816–6826.

Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217.

Novaes E, Osorio L, Drost DR, et al. 2009. Quantitative genetic analysis of biomass and wood chemistry of *Populus* under different nitrogen levels. *The New phytologist* **182**, 878–90.

Oinonen C, Rouvinen J. 2000. Structural comparison of Ntn-hydrolases. *Protein science* **9**, 2329–2337.

Oza VP, Trivedi SD, Parmar PP, Subramanian RB. 2009. *Withania somnifera* (Ashwagandha): A novel source of L-asparaginase. *Journal of Integrative Plant Biology* **51**, 201–206.

Pacheco-Villalobos D. 2008. Aproximación genómica al estudio de la formación de la madera en los pinos. Universidad de Málaga, Spain.

Pandurangan S, Pajak A, Molnar SJ, Cober ER, Dhaubhadel S, Hernández-Sebasti C, Kaiser WM, Nelson RL, Huber SC, Marsolais F. 2012. Relationship between asparagine metabolism and protein concentration in soybean seed. *Journal of Experimental Botany* **63**, 3173–3184.

Pascual MB, Cánovas FM, Ávila C. 2015. The NAC transcription factor family in maritime pine (*Pinus Pinaster*): molecular regulation of two genes involved in stress responses. *BMC Plant Biology*, 1–15.

Pauwels L, Morreel K, De Witte E, Lammertyn F, Van Montagu M, Boerjan W, Inzé D, Goossens A. 2008. Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured *Arabidopsis* cells. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 1380–1385.

Pedranzani H, Sierra-De-Grado R, Vigliocco A, Miersch O, Abdala G. 2007. Cold and water stresses produce changes in endogenous jasmonates in two populations of *Pinus pinaster* Ait. *Plant Growth Regulation* **52**, 111–116.

Roberts J. 1976. Purification and properties of a highly potent antitumor glutaminase-asparaginase from *Pseudomonas* 7A. *J Biol Chem* **251**, 2119–2123.

Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, van den hoff MJB, Moorman AFM. 2009. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Research* **37**.

Ruzicka K, Ursache R, Hejátko J, Helariutta Y. 2015. Xylem development – from the cradle to the grave. *New Phytologist* **207**, 519–535.

Sakakibara H, Watanabe M, Hases T, Sugiyama T. 1991. Molecular cloning and characterization of complementary DNA encoding for ferredoxin-dependent glutamate synthase in maize leaf. *Journal of Biological Chemistry* **266**, 2028–2035.

Salvador L, Alía R, Agúndez D, Gil L. 2000. Genetic variation and migration pathways of maritime pine (*Pinus pinaster* Ait) in the Iberian peninsula. *TAG Theoretical and Applied Genetics* **100**, 89–95.

Sanchez-Gomez D, Majada J, Alía R, Feito I, Aranda I, Sánchez-Gómez D. 2010. Intraspecific variation in growth and allocation patterns in seedlings of *Pinus pinaster* Ait. submitted to contrasting watering regimes: can water availability explain regional variation? *Annals of Forest Science* **67**, 505–513.

Sannigrahi P, Ragauskas AJ. 2010. Poplar as a feedstock for biofuels: A review of compositional characteristics. *Biofuels, Bioproducts and Biorefining* **4**, 209–226.

Scheible W-R, Morcuende R, Czechowski T, Fritz C, Osuna D, Palacios-Rojas N, Schindelasch D, Thimm O, Udvardi MK, Stitt M. 2004. Genome-wide reprogramming of primary and secondary metabolism, protein synthesis, cellular growth processes, and the regulatory infrastructure of Arabidopsis in response to nitrogen. *Plant Physiology* **136**, 2483–2499.

Shvidenko A. 2011. Changing world, boreal forests and IBFRA. *Boreal Forests in a Changing World: Challenges and Needs for Action. Proceedings of the international conference IBFRA, Krasnoyarsk, Russia*, 8–12.

Sieciechowicz K, Ireland RJ, Joy KW. 1985. Diurnal variation of asparaginase in developing pea leaves. *Plant Physiology* **77**, 506–508.

Sieciechowicz KA, Joy KW, Ireland RJ. 1988. The metabolism of asparagine in plants. *Phytochemistry* **27**, 663–671.

Sieciechowicz KA, Joy KW, Ireland RJ. 1989. Effect of methionine sulfoximine on asparaginase activity and ammonium levels in pea leaves. *Plant Physiol.* **89**, 192–196.

Sievers F, Wilm A, Dineen D, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* **7**, 539.

Slocum RD. 2005. Genes, enzymes and regulation of arginine biosynthesis in plants. *Plant physiology and Biochemistry* **43**, 729–745.

Sodek L. 1980. Distribution and properties of a potassium-dependent asparaginase isolated from developing seeds of *Pisum sativum* and other plants. *Plant physiology* **65**, 22–26.

Sodek L, Lea PJ. 1993. Asparaginase from the testa of developing lupin and pea seeds. *Phytochemistry* **34**, 51–56.

Stasolla C, Katahira R, Thorpe T a, Ashihara H. 2003. Purine and pyrimidine nucleotide metabolism in higher plants. *Journal of plant physiology* **160**, 1271–1295.

Stone SL, Gifford DJ. 1997. Structural and biochemical changes in loblolly pine (*Pinus taeda* L.) seeds during germination and early-seedling growth. I. Storage protein reserves. *International Journal of Plant Science* **158**, 727–737.

Streeter JG. 1977. Asparaginase and asparagine transaminase in soybean leaves and root nodules. *Plant physiology* **60**, 235–239.

Su Y, Karamitros CS, Nomme J, McSorley T, Konrad M, Lavie A. 2013. Free glycine accelerates the autoprolytic activation of human asparaginase. *Chemistry and Biology* **20**, 533–540.

Suárez MF, Avila C, Gallardo F, Cantón FR, García-Gutiérrez A, Claros MG, Cánovas FM. 2002. Molecular and enzymatic analysis of ammonium assimilation in woody plants. *Journal of Experimental Botany* **53**, 891–904.

Swain a L, Jaskólski M, Housset D, Rao JK, Wlodawer A. 1993. Crystal structure of *Escherichia coli* L-asparaginase, an enzyme used in cancer therapy. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 1474–1478.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729.

Todd CD, Cooke JEK, Mullen RT, Gifford DJ. 2001. Regulation of loblolly pine (*Pinus taeda* L.) arginase in developing seedling tissue during germination and post-germinative growth. *Plant Molecular Biology* **45**, 555–565.

Tomar R, Garg DK, Mishra R, Thakur AK, Kundu B. 2013. N-terminal domain of *Pyrococcus furiosus* L-asparaginase functions as a non-specific, stable, molecular chaperone. *FEBS Journal* **280**, 2688–2699.

Tripathy JN, Hirasawa M, Sutton RB, Dasgupta A, Vaidyanathan N, Zabet-Moghaddam M, Florencio FJ, Srivastava AP, Knaff DB. 2015. A loop unique to ferredoxin-dependent glutamate synthases is not absolutely essential for ferredoxin-dependent catalytic activity. *Photosynthesis Research* **123**, 129–139.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology* **3**, RESEARCH0034.

Villalobos DP, Díaz-Moreno SM, Said E-SS, Cañas R a, Osuna D, Van Kerckhoven SHE, Bautista R, Claros MG, Cánovas FM, Cantón FR. 2012. Reprogramming of gene expression during compression wood formation in pine: coordinated modulation of S-adenosylmethionine, lignin and lignan related genes. *BMC plant biology* **12**, 100.

Wang M, Kurland CG, Caetano-Anollés G. 2011. Reductive evolution of proteomes and protein structures. *Proceedings of the National Academy of Sciences* **108**, 11954–11958.

Willis RC, Woolfolk C a. 1974. Asparagine utilization in *Escherichia coli*. *Journal of bacteriology* **118**, 231–41.

Wu CS, Chaw SM. 2014. Highly rearranged and size-variable chloroplast genomes in conifers II clade (cupressophytes): Evolution towards shorter intergenic spacers. *Plant Biotechnology Journal* **12**, 344–353.

Xu Q, Buckley D, Guan C, Guo HC. 1999. Structural insights into the mechanism of intramolecular proteolysis. *Cell* **98**, 651–661.

Yang Y, Costa A, Leonhardt N, Siegel RS, Schroeder JI. 2008. Isolation of a strong *Arabidopsis* guard cell promoter and its potential as a research tool. *Plant methods* **4**, 6.

Zuckerkindl E, Pauling L. 1965. Molecules as documents of evolutionary history. *Journal of theoretical biology* **8**, 357–366.

8. ANEXOS

Suplemento 1: Referencias de las bases de datos de secuencias primarias de ASPG usadas en el análisis filogenético

<u>Abreviatura</u>	<u>Especie</u>	<u>Base de datos</u>	<u>Referencia</u>
Especies no plantas			
EcoAIII	<i>Escherichia coli</i>	GenBank	EFO58337
hASNase3	<i>Homo sapiens</i>	GenBank	AAM28434
Briofitas			
PhPaten1	<i>Physcomitrella patens</i>	Phytozome 11.0 (JGI)	Pp3c14_14930
PhPaten2	<i>Physcomitrella patens</i>	Phytozome 11.0 (JGI)	Pp3c8_21770
PhPaten3	<i>Physcomitrella patens</i>	Phytozome 11.0 (JGI)	Pp3c23_6900
SpFalla1	<i>Sphagnum fallax</i>	Phytozome 11.0 (JGI)	Sphfalx0119s0066
SpFalla2	<i>Sphagnum fallax</i>	Phytozome 11.0 (JGI)	Sphfalx0119s0065
Licofitas			
SeMoell	<i>Selaginella moellendorffii</i>	UniProtKB	D8T3N6
Gimnospermas			
PiSylve	<i>Pinus sylvestris</i>	GenBank	CAK22360
PiTaeda	<i>Pinus taeda</i>	Congenie.org	PITA_000017050
Angiospermas monocotiledóneas			
BrDista1	<i>Brachypodium distachyon</i>	Phytozome 11.0 (JGI)	Bradi5g24227
BrDista2	<i>Brachypodium distachyon</i>	Phytozome 11.0 (JGI)	Bradi1g15126
BrDista3	<i>Brachypodium distachyon</i>	Phytozome 11.0 (JGI)	Bradi1g15130
OrSativ1	<i>Oryza sativa</i>	Phytozome 11.0 (JGI)	LOC_Os04g55710
OrSativ2	<i>Oryza sativa</i>	Phytozome 11.0 (JGI)	LOC_Os03g40070
PaHalli1	<i>Panicum hallii</i>	Phytozome 11.0 (JGI)	Pahal.G02572
PaHalli2	<i>Panicum hallii</i>	Phytozome 11.0 (JGI)	Pahal.I02619
PaHalli3	<i>Panicum hallii</i>	Phytozome 11.0 (JGI)	Pahal.I02618

Angiospermas dicotiledóneas

ArThali1	<i>Arabidopsis thaliana</i>	TAIR (www.arabidopsis.org)	AT3G16150
ArThali2	<i>Arabidopsis thaliana</i>	TAIR (www.arabidopsis.org)	AT5G08100
BrRapa1	<i>Brassica rapa</i>	Phytozome 11.0 (JGI)	Brara.A03076
BrRapa2	<i>Brassica rapa</i>	Phytozome 11.0 (JGI)	Brara.E02537
BrRapa3	<i>Brassica rapa</i>	Phytozome 11.0 (JGI)	Brara.J02454
BrRapa4	<i>Brassica rapa</i>	Phytozome 11.0 (JGI)	Brara.J02455
GIMax1	<i>Glycine max</i>	UniProtKB	I1MXK5
GIMax2	<i>Glycine max</i>	UniProtKB	I1M8R0
GIMax3	<i>Glycine max</i>	UniProtKB	I1K824
GIMax4	<i>Glycine max</i>	UniProtKB	I1JTM6
GIMax5	<i>Glycine max</i>	UniProtKB	C6TER3
GoRaimo1	<i>Gossypium raimondii</i>	Phytozome 11.0 (JGI)	Gorai.005G101000
GoRaimo2	<i>Gossypium raimondii</i>	Phytozome 11.0 (JGI)	Gorai.009G276200
GoRaimo3	<i>Gossypium raimondii</i>	Phytozome 11.0 (JGI)	Gorai.010G023000
GoRaimo4	<i>Gossypium raimondii</i>	Phytozome 11.0 (JGI)	Gorai.009G215700
GoRaimo5	<i>Gossypium raimondii</i>	Phytozome 11.0 (JGI)	Gorai.008G200400
GoRaimo6	<i>Gossypium raimondii</i>	Phytozome 11.0 (JGI)	Gorai.007G040100
MeTrunc1	<i>Medicago truncatula</i>	UniProtKB	Q2HTR7
MeTrunc2	<i>Medicago truncatula</i>	UniProtKB	G7I4Z5
MeTrunc3	<i>Medicago truncatula</i>	UniProtKB	G7J705
MeTrunc4	<i>Medicago truncatula</i>	UniProtKB	A0A072URR5
MeTrunc5	<i>Medicago truncatula</i>	UniProtKB	A0A072USJ6
PhVulga1	<i>Phaseolus vulgaris</i>	UniProtKB	V7CU13
PhVulga2	<i>Phaseolus vulgaris</i>	UniProtKB	V7AST5
PhVulga3	<i>Phaseolus vulgaris</i>	UniProtKB	V7CAP3
PoTrich1	<i>Populus trichocarpa</i>	Phytozome 11.0 (JGI)	Potri.014G022900
PoTrich2	<i>Populus trichocarpa</i>	Phytozome 11.0 (JGI)	Potri.002G122900
PoTrich3	<i>Populus trichocarpa</i>	Phytozome 11.0 (JGI)	Potri.012G063400
SoLycop1	<i>Solanum lycopersicum</i>	UniProtKB	K4BUT1
SoLycop2	<i>Solanum lycopersicum</i>	UniProtKB	K4C8B3
SoLycop3	<i>Solanum lycopersicum</i>	UniProtKB	K4BKY1
ViVinif1	<i>Vitis vinífera</i>	Phytozome 11.0 (JGI)	GSVIVG01009430001
ViVinif2	<i>Vitis vinífera</i>	Phytozome 11.0 (JGI)	GSVIVG01008150001

Suplemento 2: Cebadores usados en las reacciones de PCR

Referencia	Secuencia nucleotídica
A1-1	GGGTGCATCCTACTGCCAAAT
A1-2	GCCACCCTTACCATCTTCCAA
A1-3	GTGGTTGTTCTGTGGGAAGA
A1-4	CTAGTTAATTAACATGGGATGGGCCATAGCT
A1-5	AGAGGCGGCCGCTACTGCCAAATGCC
A1-6	GGAATGTTCCCGGAGACTGTGGGTTGTGTTG
A1-7	CAACACAACCCACAGTCTCCGGGAACATTCC
A1-8	CCTTCAGAATGGGGAAACTGTGGGTTGTGTTG
A1-9	CAACACAACCCACAGTTTCCCCATTCTGAAGG
A1-10	GTTTGACTACCGTCTCCAGAAACTGTGGGTTGTGTTG
A1-11	CAACACAACCCACAGTTTCTGGGAGACGGTAGTCAAAC
A1-12	CATCGCGGCCGCTTAGTGATGGTGATGGTGATGCGCATAGTCAGGAACAT CGTATGGGTACTGCCAAATGCCTACTTC
A1-13	GGTGAAGGTGAAGCCATAATCC
A1-14	TGCAGCAACGTCTCCATTAC
A1-15	AAAAAGCAGGCTTCATGGGATGGGCCATAGCT
A1-16	AGAAAGCTGGGTCCTACTGCCAAATGCCTACTTCAAA
A2-1	GGAAGGCGAGAGGCAG
A2-2	GCATACCACATAACAATCCCAAG
A2-3	GAGGTTAATTAACATGGGGTGGGCAATTGCCTTG
A2-4	AGAAGCGGCCGCTTAGTGATGGTGATGGTGATGCGCATAGTCAGGAACAT CGTATGGGTATGCCAATATTCCCACTTC
A2-5	GAAGCAGCTAATTATGTGATAGGAC
A3-1	GAAGGGGAGGGACGAG
A3-2	ATGCATGTTAAGACATCAGTT
A3-3	GAGCTTAATTAACATGGGGTGGGCACTTGCGTTG
A3-4	CAGTGCGGCCGCTTAGTGATGGTGATGGTGATGCGCATAGTCAGGAACAT CGTATGGGTATGCCAATATTCCCACTTC
A3-5	GCGGCTAATTATGTTATGGGAA
A3-6	CATCAGTTCAAATCCAAATACATATTAAG
AP-1	GTTGATCCATGCTTGCCAAT
AP-2	GACCAGAACTGCTCCACA
AP-3	CATTAATTGGATTTGGATTGCTCACCAGG
AP-4	CAATGCAAGAGCAGGTCCTTGGCTTC
AP-5	TGTCAAGCTTTCCAATTATTGTCTCTA
AP-6	CAATAAGCTTTCCATATTAACAATTCTAA
AP-7	AACAAAGCTTGAACAACCATGAATTAC

AP-8	CAGCGGATCCAAAGGGAAGTGGA
AP-9	CAGCGAATACGTCCATTTCC
AP-10	GACAATGCAAGAGCAGGTC
AP-11	CTGCCTATCACGTTGTGT
AP-12	AAAGTGTGCCTAGGATATGA
AP-13	AAACATCTAGGGAACAACCA
OA-1	CCTGAACCATGTGATCGTTG
OA-2	CTAAGGCAATGTGGGCATTT
GS-1	GGAACGTGAAGGCAAAGG
GS-2	CACTTCAATCGTATCGGCATC
attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCT
attB2	GGGGACCACTTTGTACAAGAAAGCTGGGT



Characterization of Three L-Asparaginases from Maritime Pine (*Pinus pinaster* Ait.)

Sonia H. Van Kerckhoven, Fernando N. de la Torre, Rafael A. Cañas, Concepción Avila, Francisco R. Cantón* and Francisco M. Cánovas

Departamento de Biología Molecular y Bioquímica, Facultad de Ciencias, Universidad de Málaga, Málaga, Spain

OPEN ACCESS

Edited by:

Wanchai De-Eknamkul,
Chulalongkorn University, Thailand

Reviewed by:

Nitima Suttipanta,
Ubon Ratchathani University, Thailand
Takahiro Mori,
ETH Zurich, Switzerland

*Correspondence:

Francisco R. Cantón
frcanton@uma.es

Specialty section:

This article was submitted to
Plant Metabolism
and Chemodiversity,
a section of the journal
Frontiers in Plant Science

Received: 17 March 2017

Accepted: 06 June 2017

Published: 23 June 2017

Citation:

Van Kerckhoven SH, de la Torre FN,
Cañas RA, Avila C, Cantón FR and
Cánovas FM (2017) Characterization
of Three L-Asparaginases from
Maritime Pine (*Pinus pinaster* Ait.).
Front. Plant Sci. 8:1075.
doi: 10.3389/fpls.2017.01075

Asparaginases (ASPG, EC 3.5.1.1) catalyze the hydrolysis of the amide group of L-asparagine producing L-aspartate and ammonium. Three ASPG, PpASPG1, PpASPG2, and PpASPG3, have been identified in the transcriptome of maritime pine (*Pinus pinaster* Ait.) that were transiently expressed in *Nicotiana benthamiana* by agroinfection. The three recombinant proteins were processed *in planta* to active enzymes and it was found that all mature forms exhibited double activity asparaginase/isoaspartyl dipeptidase but only PpASPG1 was able to catalyze efficiently L-asparagine hydrolysis. PpASPG1 contains a variable region of 77 amino acids that is critical for proteolytic processing of the precursor and is retained in the mature enzyme. Furthermore, the functional analysis of deletion mutants demonstrated that this protein fragment is required for specific recognition of the substrate and favors enzyme stability. Potassium has a limited effect on the activation of maritime pine ASPG what is consistent with the lack of a critical residue essential for interaction of cation. Taken together, the results presented here highlight the specific features of ASPG from conifers when compared to the enzymes from angiosperms.

Keywords: conifers, maritime pine, asparagine, aspartate, nitrogen metabolism, proteolytic processing

INTRODUCTION

Asparaginases (ASPG, EC 3.5.1.1) catalyze the hydrolysis of the amide group of L-asparagine (Asn) producing L-aspartate (Asp) and ammonium. Some ASPG are also able to hydrolyze modified substrates such as glycosylated L-Asn or β -aspartyl dipeptides, resulting in the release of L-Asp. In 1980, Sodek reported the presence of two different ASPG activities in plants, dependent or independent of potassium ion (K^+) for their activity. Bruneau et al. (2006) confirmed the co-existence of two ASPG in *Arabidopsis thaliana* and it was initially observed that the K^+ -dependent type is strictly specific for Asn whereas the K^+ -independent ASPG can also accept isoaspartyl dipeptides as substrates. Similar results were found in *Lotus japonicus* (Credali et al., 2011), but recent work on the *Arabidopsis* ASPG also reports a double activity asparaginase/isoaspartyl dipeptidase for the K^+ -dependent type (Gabriel et al., 2012). Crystallographic structure analysis of a K^+ -dependent ASPG from *Phaseolus vulgaris* showed a catalytic pocket big enough to host substrates larger than Asn, but it has been proposed that substrate specificity is probably determined by the nature of the amino acids in the vicinity of the substrate side chain

(Bejger et al., 2014). The existence of these two subgroups within plants suggests a functional specialization for the isoforms.

Plant ASPG belong to the superfamily of Ntn (N-terminal nucleophile) hydrolases which include enzymes with large diversity in catalytic activity (Galperin and Koonin, 2012) and different oligomeric forms (Michalska and Jaskólski, 2006). Members of this protein family share a typical fold which incorporates a nucleophilic residue (serine, threonine, or cysteine) at the N-terminal incorporated in a β -sheet, providing the capacity for a nucleophilic attack and the possibility of autocatalytic processing (Brannigan et al., 1995). Plant ASPG, as well as aspartylglucosaminidase and taspase1, occur in a heterotetrameric $(\alpha\beta)_2$ structure which results from the combination of two autocatalytically cleaved precursors into their α and β subunits (Michalska and Jaskólski, 2006).

The autocatalytic activation reaction is a posttranslational modification *in cis* which, apart from the nucleophile, requires a precise alignment of several functional groups (Guan et al., 1996; Xu et al., 1999). When comparing protein sequences of ASPG isoforms and homologs in plants, high diversity is noticed in the region preceding the catalytic nucleophile, both in sequence as well as in length. This region, often referred to as a variable loop (VL), presents a lack of electron density in X-ray diffraction data of the protein crystals and therefore probably does not have a defined secondary structure (Michalska et al., 2006; Bejger et al., 2014). This has made it difficult to postulate a hypothesis about its function based on structural aspects. However, there are evidences that this region may have a role in autocatalytic processing and substrate specificity. In particular, reciprocal exchange of the C-terminal end of the α subunit between *Arabidopsis* K^+ -dependent and K^+ -independent enzymes affected the affinity for Asn and β -Asp-His, and deletions in this region impacted self-cleavage processing and catalytic activity (Gabriel et al., 2012). It has been postulated that the autocatalytic activation mechanism and amino acid residues involved differ between the ASPG of different species and possibly a certain flexibility in the region that precedes the nucleophile is necessary (Michalska et al., 2008; Nomme et al., 2012). Recombinant plant ASPG expressed in *E. coli* have been found to mature more slowly but the autocatalytic activation process continues *in vitro* (Michalska et al., 2006; Cantor et al., 2009; Nomme et al., 2012). *Pinus sylvestris* ASPG, however, did not present autoproteolytic capacity *in vitro* but the addition of root extracts from young trees resulted in enzyme activation (Cañas et al., 2007). These findings suggest a different maturation mechanism for these enzymes in gymnosperms, which could require the assistance of an external factor. The primary structure of the *P. sylvestris* ASPG, when compared to that of angiosperms, is about 50 amino acids larger due to a non-conserved sequence located in the VL.

Asn is a metabolite used by plants as a vehicle for storage and transport of nitrogen (N) (Borek and Jaskólski, 2001) and, therefore, as a N donor in sink tissues through ASPG activity (Sieciechowicz et al., 1988; Gaufichon et al., 2013). The use of this N donor differs temporally and between species, organs, and tissues (Lea et al., 2007). On the other hand, Asp is an immediate precursor for the biosynthesis

of other amino acids of the Asp family (lysine, threonine, methionine) (Azevedo et al., 2006) which can also be substrates for further metabolic reactions that produce isoleucine, glycine, and serine (Hildebrandt et al., 2015). ASPG activity therefore could contribute in providing substrates for several metabolic pathways involved in development and growth. In the hypocotyls of pine seedlings, Asn seems to be the main N storage molecule during early pine development and a temporal and spatial correlation between ASPG expression and appearance of the vascular system in pine suggests a function for this enzyme in N mobilization to sustain the development of this structure (Cañas et al., 2007). Thus, asparaginase appears to play an important role during development of the xylem in the hypocotyl once the pine seedlings have consumed the seed reserves. This is essential for distribution of water and nutrients from the roots and therefore for seedling survival after germination.

Pinus pinaster Aiton is one of the most common conifer species in the occidental Mediterranean basin where it is of great ecological importance, but also highly valued for the production of wood and resin. Pine trees are tolerant to diverse types of stress and are able to grow on poor soil and problematic habitats (Lev-Yadun and Sederoff, 2000). Compared to other pine species, *P. pinaster* populations present a high rate of genetic variability, polymorphisms and heterozygosity (Salvador et al., 2000) which contributes to its value as a model species and facilitates identification of suitable genotypes for domestication. In this work, we present the biochemical characterization of pine asparaginases and new insights into the function of the VL sequence in conifers.

MATERIALS AND METHODS

Biological Material

Escherichia coli strains DH5 α and BL21-AI (Invitrogen Corporation, Carlsbad, CA, United States) were used for plasmid propagation and expression of the protein in a prokaryotic system, respectively. For protein production in *Nicotiana benthamiana*, *Agrobacterium tumefaciens* strain C58C1 was applied. *P. pinaster* seeds from Sierra Bermeja provenance (Estepona, Spain) were germinated and grown as previously described (Cañas et al., 2007).

RNA Extraction and cDNA Synthesis

Tissue samples were ground under liquid nitrogen with a mortar and pestle. Total RNA was extracted from 100 mg ground powder as previously described (Canales et al., 2012). The integrity of RNA in the samples was verified by agarose gel electrophoresis.

To obtain the full-length cDNA for the *PpASPG1* gene, cDNA was generated following the indications of the “FirstChoice RLM RACE” kit (Applied Biosystems/Ambion, Austin, TX, United States) using 10 μ g of total RNA extracted from xylem as a template. For amplification of the *PpASPG2* and *PpASPG3* cDNAs 1 μ g of total RNA extracted from radicles of seedlings was used to synthesize single strand cDNA using iScript cDNA Synthesis kit (Bio-Rad, Hercules, CA, United States).

PCR Amplification and Cloning of ASPG cDNAs

All PCR reactions were carried out with iProof High-Fidelity DNA Polymerase (Bio-Rad, Hercules, CA, United States). Reverse primers A1-1 and A1-2 (see Supplementary Table S1) were designed on a previously known EST sequence (Villalobos et al., 2012) to amplify the 5'UTR region of the *PpASPG1* gene using 5'RACE cDNA template. PCR reactions were performed under the following conditions: initial denaturalization for 2 min at 95°C, 30 cycles of 30 s at 95°C, 30 s at 60°C and 1 min 15 s at 72°C with a final extension of 10 min at 72°C. The amplified DNA fragment was blunt end ligated in the vector pBluescript SK+ (Agilent Technologies, Santa Clara, CA, United States) and propagated by transforming *E. coli*. Inserts were verified by sequencing with a Beckman Coulter CEQ8000 sequencer. Forward primer A1-3 was designed over the 5'UTR region to obtain the full-length cDNA using the 3'RACE cDNA as a template. The PCR product was also ligated into vector pBSK and transformed into *E. coli* strain DH5 α competent cells for propagation and sequencing. The final plasmid preparation was used as a template for constructions in expression vectors.

Primers A1-4 and A1-5 were used to amplify the complete ORF of *PpASPG1* without any tags, incorporating a unique restriction site for *PacI* in the N-terminal region and a *NotI* site in the C-terminal region. cDNAs were amplified using primers A2-1 and A2-2 for *PpASPG2* and primers A3-1 and A3-2 for *PpASPG3* and cDNA of young pine plantlets as a template. Primers were designed on the UTR regions of unigenes 15077 and 36994 available at the Sustainpine V.3.0 database (Canales et al., 2014). A second PCR reaction was carried out to amplify the ORF and to incorporate the restriction sites for *PacI* and *NotI* to facilitate the cloning in the expression vector pJL-TRBO, as well as an HA and poly-His tag at the C-terminal, using primers A2-3 and A2-4 for *PpASPG2* and primers A3-3 and A3-4 for *PpASPG3*.

Generation of Deletion Constructs by PCR

To generate *PpASPG1* cDNAs with different deletions in the internal variable region, sets of two deletion primers with complementary sequences were designed. Half of the sequence of each primer was derived from 5'-flanking region of the sequence to be eliminated and the other half from the 3'-flanking region. Deletion primers A1-6 and A1-7 were used for D1, A1-8 and A1-9 for D2 and A1-10 and A1-11 for D3 construction (see Supplementary Table S1 for primer sequences). For every deletion two rounds of PCR were carried out. In the first round, two PCR reactions were performed, each one with one primer of the corresponding pair of complementary deletion primers and a second primer derived from one end of the complete ORF sequence (A1-4 forward and A1-12 reverse primers). This procedure was used to separately amplify the fragments covering the sequence for the α and β subunit (for example, for deletion D1, A1-6 was combined with A1-12 to produce the β -subunit encoding fragment and A1-7 with A1-4 to produce the α -subunit encoding fragment). Consequently, the 3'-end of the amplified sequence for the α subunit and the 5'-end of the sequence for

the β subunit had the same sequence, corresponding with the sequence of the primers A1-6 and A1-7 for D1, A1-8 and A1-9 for D2 and A1-10 and A1-11 for D3. For each deletion, the products of the first PCR reactions were then combined as template in a second round of PCR, and amplified with the primers A1-4 and A1-12. After denaturing and annealing in the first PCR cycle, sense DNA strands derived from the sequence for the α subunit can be annealed with the antisense DNA strand derived from the sequence for the β subunit through the complementary region introduced by the deletion primers, and each strand will act as a primer for the synthesis of the complementary strand by the DNA polymerase. This new product consisting in the precursor coding sequence, but with the selected internal sequence deleted, was amplified in the subsequent PCR cycles with the A1-4 and the A-12 primers. The A1-4 forward primer included a *PacI* restriction site and the A-12 reverse primer a *NotI* restriction site to facilitate the cloning in the expression vector pJL-TRBO. The reverse primer A1-12 also includes a C-terminal HA and poly-His tag sequence. For comparison, the complete sequence of the *PpASPG1* ORF was also amplified with these primers to include the tag.

Construction of Recombinant Expression Vectors for *In planta* Expression

For *in planta* expression, the different PCR products derived from the ORFs described above were digested with restriction enzymes *PacI* and *NotI* Fast Digest (Thermo Fisher Scientific Inc., Waltham, MA, United States) and ligated into the pJL-TRBO vector (Lindbo, 2007) previously digested with the same restriction enzymes.

In planta Protein Expression, Extraction, and Western Blot Analysis

Agrobacterium tumefaciens strain C58C1 was transformed with the expression vector pJL-TRBO by electroporation. Overnight grown cells harvested by centrifugation were used to infiltrate *N. benthamiana* leaves with a syringe by applying slight pressure on a small incision on the abaxial side of the leaves. Five days later the infiltrated areas of the *N. benthamiana* leaves were homogenized in liquid nitrogen. Extraction buffer [50 mM Hepes pH 7.8, 50 mM KCl, 5 mM EDTA, 10% glycerol (v/v), 10 mM 2-mercaptoethanol and 1 mM PMSF] was added in proportion 2:1 (v/w). Crude extracts were obtained by centrifugation at $18,660 \times g$ at 4°C for 20 min. Successful expression *in planta* was confirmed by western blot analysis and asparaginase activity of crude extracts. A negative control consisting in crude extracts from *N. benthamiana* leaves infiltrated with *A. tumefaciens* transformed with the empty pJL-TRBO vector was always included as a reference.

Laemmli buffer (Laemmli, 1970) as well as 0.5 mM 2-mercaptoethanol were added to the crude extracts and proteins resolved on 12.5% (w/v) polyacrylamide gels. The gels were transferred onto nitrocellulose membranes (Whatman GmbH, Dassel, Germany) and the presence of ASPG polypeptides was immunorevealed as described previously (Cañas et al., 2007), using either the antiserum raised against pine ASPG or an

antibody against the HA tag. Detection of immunocomplexes was carried out by chemiluminescence using the SuperSignal West Pico Chemiluminescent Substrate system (Thermo Fisher Scientific Inc., Waltham, MA EE.UU., United States). Protein levels were determined by the Bradford procedure (Bradford, 1976).

Protein Purification of Recombinant Proteins Produced *In planta*

The PpASPG1 and PpASPG1-D3 recombinant proteins were purified from *N. benthamiana* leaf extracts mostly as described by Sodek and Lea (1993) with the following exceptions and specifications. Soluble proteins extracted from the infiltrated tobacco leaf areas were concentrated by the gradual addition of $(\text{NH}_4)_2\text{SO}_4$ up to 80% saturation. The pellet was resuspended in extraction buffer, ethanol was added up to a concentration of 60% (v/v), and incubated on ice for 1 h. After centrifugation at $2,000 \times g$ for 20 min at 4°C insoluble proteins were discarded and ASPG was recovered in the supernatant. The mixture was extensively dialyzed in extraction buffer to remove the ethanol, after which the protein was concentrated by reverse dialysis in sucrose. The concentrate was loaded onto a DEAE-Sephacel (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) 50 mL column. Elution was performed with a gradient formed of 50 mM KCl and 500 mM KCl at a flow rate of 42 mL/h. The fractions with ASPG activity were pooled and concentrated by reverse dialysis in sucrose before subjection to size-exclusion gel chromatography on a Sephacryl S-300 column (Sigma-Aldrich Co. LLC, St. Louis, MO, United States) (93 cm \times 2 cm) at a flow rate 16 mL/h. Fractions with ASPG activity were pooled, flash-frozen in liquid nitrogen and stored at -80°C until further use.

PpASPG2 and PpASPG3 recombinant proteins were purified from the leaf extracts applying Protino® Ni-TED 2000 columns (MACHEREY-NAGEL GmbH & Co KG, Düren, Germany), following manual instructions.

Enzyme Activity Assays and Biochemical Characterization

The enzymatic activity of recombinant ASPG was determined by the rate of NH_4^+ production during the hydrolysis of Asn. Aliquots of 1 μL of the purified protein were incubated during 30 min at 37°C in a volume of 15 μL activity buffer (50 mM HEPES pH 7.8, 50 mM KCl and 20 mM asparagine). The reactions were stopped by heating the samples for 2 min in a bath with boiling water followed by centrifugation at $12,400 \times g$ for 2 min. Free ammonia quantification was determined in the supernatant by the Berthelot phenol-hypochlorite method (Loan et al., 2013). After 30 min at room temperature absorbance at 635 nm was measured. Free ammonia was calculated using a standard curve obtained with different concentrations of ammonium sulfate. One unit of enzyme activity (1U) was defined as the amount of enzyme required to release 1 μmole of NH_4^+ per min.

To determine optimal temperature, samples were incubated in a Biometra Tgradient thermocycler (Biometra GmbH, Göttingen, Germany) with a 5°C interval temperature gradient ranging

from 5 to 60°C . Enzymatic activity was determined by the rate of NH_4^+ production, as explained above. The activation energy was calculated by applying Arrhenius equation to the plotted data. The thermal stability was evaluated by incubating aliquots of the enzyme at temperatures ranging from 40 to 70°C without the presence of the Asn substrate, and determining enzymatic activity by the rate of NH_4^+ production. Relative activity for two different substrates (L-Asn and β -L-Asp-L-Ala) was determined by measuring the release of aspartate as described by Möllering (1985). Reactions were performed in microplates in a volume of 180 μL containing 50 mM HEPES pH 7.8, 50 mM KCl, 3.5 mM α -ketoglutarate, 0.6 mM NADH, 5 U each of glutamate-oxaloacetate transaminase (GOT) and malate dehydrogenase (MDH), and the substrate at the corresponding concentration. NADH oxidation was monitored using a “PowerWave HT Microplate Spectrophotometer” (Biotek Instruments, Winooski, VT, United States). To determine K_m and V_{\max} values the assay based on measuring the release of aspartate described above was used. The concentrations of the substrates present in the reaction ranged from 0.2 to 10 times the value of K_m . The experimental data were fitted to a Michaelis-Menten equation and the kinetic behavior was calculated by the Hanes-Woolf method ($r^2 > 0.99$).

Construction of Recombinant Vectors for PpASPG1 Expression in *E. coli*

Primers A1-13 and A1-14 were used to amplify the region spanning the ORF of PpASPG1 or D3 variant using the plasmids described for *in planta* expression as templates. A secondary PCR reaction was performed with primers attB1 and attB2 for the introduction in Gateway vectors (Invitrogen Corporation, Carlsbad, CA, United States). The fragment was inserted in vector pDONR207 by recombination and transferred to the expression vector pDEST17. The expression vector was transferred to *E. coli* strain DH5 α by electroporation for propagation and verification.

Protein Expression in *E. coli* and Purification

The *E. coli* strain BL21-AI was transformed with the pDEST17. Protein expression was induced by adding 0.2% (w/v) arabinose.

Recombinant proteins were purified from *E. coli* inclusion bodies in denaturing conditions, applying Protino® Ni-TED 2000 columns (MACHEREY-NAGEL GmbH & Co KG, Düren, Germany), following manual instructions. Protein samples were flash-frozen in liquid nitrogen and stored at -80°C . The electrophoretically separated proteins were visualized by silver stain following the recommendations of the “Silver stain Plus” kit (Bio-Rad, Hercules, CA, United States).

MALDI-TOF Mass Spectrometry

The PpASPG1 purified protein was analyzed with a 4700 Proteomics Analyzer mass spectrometer (ABSCIEX) in linear mode to determine the molecular weight of the subunits. Subunits were also separated by SDS-PAGE and visualized by Coomassie colloidal staining. Spots were excised from the acrylamide gel and trypsin digested. Peptide mass fingerprinting

(PMF) was acquired by working in the reflector positive ion mode. Peptide sequences were confirmed by TOF-TOF PSD (post source decay) fragmentation spectra.

In Silico Analysis

Asparaginases amino acid sequences were obtained from following databases: <https://phytozome.jgi.doe.gov/>, <http://arabidopsis.org/>, <http://congenie.org/> and <http://www.scbi.uma.es/sustainpinedb/>. A list of the different sequences with the corresponding ID and the database of provenance can be found in Supplementary Table S2. Sequence alignments and phylogenetic analysis were performed using the Mega6 software (Tamura et al., 2013).

RESULTS

A search in the Sustainpine Database identified three different sequences including a full ORF encoding asparaginase in *P. pinaster* (sp_v3.0_unigene4029, sp_v3.0_unigene15077 and sp_v3.0_unigene36994), that were named *PpASPG1*, *PpASPG2* and *PpASPG3*, respectively. The phylogenetic analysis comparing the primary structure of the three proteins with plant ASPG of different species classified *PpASPG1* as a K^+ -dependent asparaginase and *PpASPG2* and *PpASPG3* as K^+ -independent ASPG (Figure 1). The *PpASPG1* sequence is 1396 nucleotides in size, including an ORF of 1128 nucleotides whereas both *PpASPG2* and *PpASPG3* sequences include a shorter ORF of 918 nucleotides. The difference in length of the ORFs is mainly the result of a higher number of codons encoding amino acids in the VL of the α -subunit C-terminus of *PpASPG1* (Supplementary Figure S1).

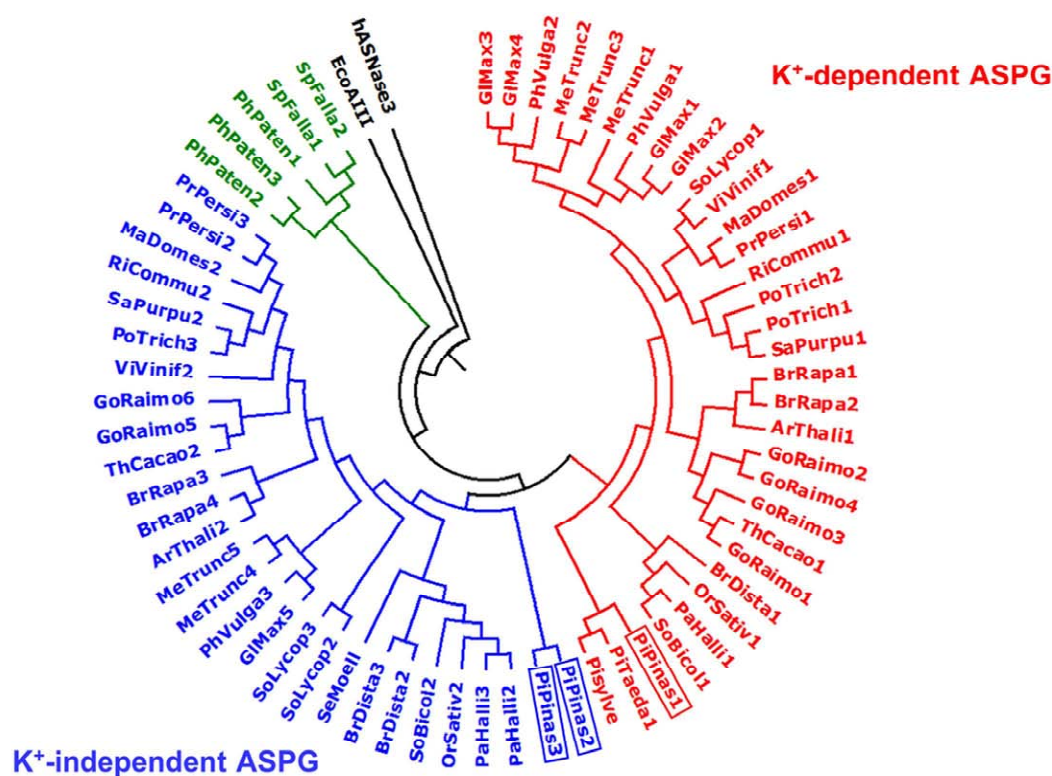
The alignment in Supplementary Figure S1 includes plant ASPG that have been biochemically characterized and determined to be K^+ -dependent or K^+ -independent enzymes. The amino acids involved in anchoring the product in the active site and the catalytic switch are absolutely conserved in all the proteins in the alignment (Supplementary Figure S1), whereas those residues determining substrate specificity show some variability but with higher conservation among K^+ -dependent enzymes. According to Beijer et al. (2014), K^+ -dependence is linked to the VMDKSPHS motif (shaded in gray and labeled as activation in Supplementary Figure S1), and it was proposed that the last serine residue is involved in the potassium-dependence mechanism. However, the *PpASPG1* protein has an apolar valine residue in the last position of this motif instead of a polar serine residue. This change is also present in the homolog proteins from *P. sylvestris* and *Pinus taeda*. As previously reported for *P. sylvestris* (Cañas et al., 2007), one significant difference between *PpASPG1* and other K^+ -dependent ASPG in plants is the number of residues that compose the VL, which is exceptionally longer in the protein from *P. pinaster* and other pine species, such as *P. sylvestris* and *P. taeda*.

The three cDNAs were transiently expressed in *N. benthamiana* leaves by agroinfection with the pJL-TRBO vector harboring each one of the three recombinant proteins as schematized in Figure 2A. All the three proteins were processed

in planta, as indicated by the size of the peptide recognized by the anti-HA-tag antibody, which corresponds to the theoretical molecular mass for the processed β subunit (Figure 2B). Nevertheless, whereas *PpASPG1* was fully processed, only part of the *PpASPG2* and *PpASPG3* precursors were processed, as shown by the remaining precursor band in the western blot analysis. The recombinant proteins produced in *N. benthamiana* were also purified and the specific activity of each protein was determined for two substrates, L-Asn and the isoaspartyl dipeptide β -L-Asp-L-Ala. *PpASPG1* showed the highest activity for L-Asn, whereas the activities of *PpASPG2* and *PpASPG3* for this substrate were only 11 and 20% of the observed value for *PpASPG1* (Figure 2C). *PpASPG3* was the enzyme with the highest activity for the substrate β -L-Asp-L-Ala, and the activities of *PpASPG1* and *PpASPG2* were 23 and 16% of the activity observed for *PpASPG3*.

The exceptionally long VL of *PpASPG1* together with the observation that the last 17 residues of the α subunit are removed from the mature recombinant EcAIII enzyme after the initial autolysis (Borek et al., 2004) prompted us to analyze the occurrence of the variable region in the mature enzyme. For pursuing this task with enough resolution to resolve the prevalence or elimination of amino acid residues in this region, we analyzed a purified preparation of recombinant and active protein by mass spectrometry. The molecular masses of both subunits were determined by MALDI-TOF mass spectrometry. As shown in Figure 3A, two peaks with masses of 26,431 and 13,325 were identified in the mass spectrum, quite close to the theoretical masses of the α and β subunits (26,535 and 13,327, respectively). Therefore, the experimental mass of the α subunit is slightly smaller than the theoretical mass. Both subunits were separated by SDS-PAGE and all bands with molecular mass in the range of 13 to 28 kDa were selected for trypsin digestion and analysis by MALDI TOF spectrometry in reflector mode, to ensure the identification of all possible variants of the processed α subunit. A total of 10 bands excised from the electrophoresis gel were analyzed. This analysis confirmed that the peptides of 26,431 and 13,325 were the α and β subunits of *PpASPG1*. The trypsin derived peptides from both subunits are shown in Figure 3B, which cover 76.6% of the α subunit and 60.3% of the β subunit. The results of mass profiles of the trypsin-derived peptides showed that the α subunit retains the complete VL in the mature protein. However, the N-terminal methionine residue is not present in the mature enzyme. Consequently, after the N-terminal methionine is removed, the theoretical molecular mass of the α subunit is 26,386 Da. The difference of 45 Da with the experimental molecular mass could be the result of a posttranslational modification. The occurrence of the complete VL in the mature *PpASPG1* protein was further confirmed by analyzing the tryptic peptide from the C-end of the α subunit by MALDI TOF/TOF tandem mass spectrometry.

The role of this exceptionally long variable region at the C-end of the α subunit in the maturation and catalytic properties of the enzyme was analyzed by producing three variants of the recombinant *PpASPG1* protein. Deletions in the ORF were generated to produce proteins lacking 14 (*PpASPG1*-D1), 42 (*PpASPG1*-D2), and 77 (*PpASPG1*-D3) residues from the



Non-plant species

EcoAIII *Escherichia coli* plant type ASPG

hASNase3 *Homo sapiens* ASPG like protein

Bryophytes

PhPatena *Physcomitrella patens* (3)

SpFallal *Sphagnum fallax* (2)

Lycophytes

SeMoell *Selaginella moellendorffii* (1)

Gymnosperms

PiPinas *Pinus pinaster* (3)

PiTaeda *Pinus taeda* (1)

PiSylve *Pinus sylvestris* (1)

Angiosperms - monocotyledons

BrDista *Brachypodium distachyon* (3)

OrSativ *Oryza sativa* (2)

PaHalli *Panicum hallii* (3)

SoBicol *Sorghum bicolor* (2)

Angiosperms - eudicotyledons

ArThali *Arabidopsis thaliana* (2)

BrRapa *Brassica rapa* (4)

GIMax *Glycine max* (5)

GoRaimo *Gossypium raimondii* (6)

MaDomes *Malus domestica* (2)

MeTrunc *Medicago truncatula* (5)

PhVulga *Phaseolus vulgaris* (3)

PoTrich *Populus trichocarpa* (3)

PrPersi *Prunus persica* (3)

RiCommu *Ricinus communis* (2)

SaPurpu *Salix purpurea* (2)

SoLycop *Solanum lycopersicum* (3)

ThCacao *Theobroma cacao* (2)

ViVinif *Vitis vinifera* (2)

FIGURE 1 | Circular dendrogram of plant-type L-asparaginases. Plant species whose genomes have been sequenced were explored for ASPG coding genes and entries with the highest similarity to pine ASPG were selected. The plant-type ASPG from *E. coli* and a human ASPG-like protein were included as outgroups. Clustering of the complete amino acid sequences was based on neighbor-joining analysis using MEGA version 6 software (Tamura et al., 2013). A bootstrap consensus tree was inferred from 1000 replicates. Red branches contain K⁺-dependent ASPG, blue branches include K⁺-independent ASPG. ASPG from bryophytes are marked in green. Below the phylogenetic tree the species included in the analysis and their corresponding abbreviations are listed. The number of different ASPG that were included for each species is mentioned in brackets. *P. pinaster* asparaginases are framed in rectangles in the dendrogram.

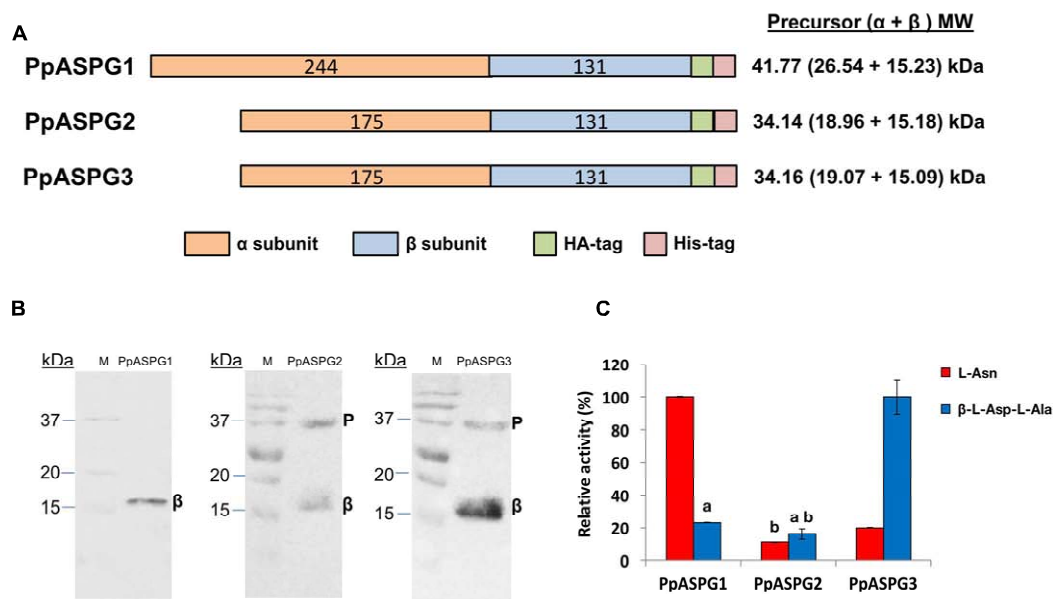


FIGURE 2 | Expression of recombinant pine ASPG in *N. benthamiana* leaves and substrate preference. **(A)** Schematic representation of the structure of the three recombinant proteins that were expressed in tobacco leaves via agroinfection. The number of amino acids is included in the colored bars and the theoretical molecular mass corresponding to the precursor and subunits that would result from correct cleavage is indicated on the right. **(B)** Western blot analysis of tobacco leaf extracts expressing the recombinant proteins. Total soluble proteins (50 μ g) from each leaf extract were separately analyzed by western blot with an antibody against the HA-tag. Bands marked with β correspond to molecular mass of the β subunit including the HA and His tags. Remaining uncleaved precursor (P) is visible in the PpASPG2 and PpASPG3 protein extracts. **(C)** Recombinant proteins were purified from the protein extracts and their relative activity was determined for two different substrates. For each substrate, catalytic activity is presented as the percentage of the highest specific activity measured at a substrate concentration of 25 mM. Values are the means of three independent replicates \pm SD. Values labeled with the same letter are not significantly different according to *t*-tests ($p < 0.01$). The remainder values are significantly different.

C-terminal of the α subunit (Figure 4A). The full-length cDNA (FL) and those encoding mutant proteins were individually expressed in *N. benthamiana* leaves, and analyzed by western blot using an antibody against the HA-tag added at the C-end of the β subunit (Figure 4B). A polypeptide with a molecular mass corresponding to the β subunit was observed for all constructions, indicating that PpASPG1-D1, PpASPG1-D2, and PpASPG1-D3 were processed and the mature protein produced. This was further confirmed by detection of asparaginase activity in crude extracts, with detection of higher levels of asparaginase activity than in control leaves transformed with the empty pJL-TRBO vector (Figure 4C). However, in the crude extracts of leaves expressing any of the three modified proteins unprocessed precursor was detected whereas no precursor polypeptide was detected in crude extracts of leaves expressing the full PpASPG1 protein. The higher amount of unprocessed precursor was observed for PpASPG1-D3, which lacked the complete variable region. This observation was consistent in different replicas of the same experiment. Additionally, in crude extracts of leaves expressing PpASPG1-D1 and PpASPG1-D2 variants, β subunit polypeptides with abnormal higher molecular masses were observed. Since the HA-tag was added at the C-terminal of the β subunit, we hypothesize that those polypeptides are the result of incorrect processing of the precursor.

To further characterize the influence of the variable region in the processing of PpASPG1 precursor we analyzed the ability

of the recombinant protein produced in *E. coli* to undergo auto-processing *in vitro*. When both PpASPG1 and PpASPG1-D3 were expressed in *E. coli* and purified by using a His-tag added at the N-end of the precursor, most protein purified from inclusion bodies accumulated unprocessed (Figure 5A). Nevertheless, some processed α and β subunits were detected for PpASPG1 but not for the PpASPG1-D3 variant (Figure 5A, time 0). Increased levels of processed α and β subunits were observed for the PpASPG1-D3 variant after 24, 48, and 72 h of incubation at room temperature but no further processing was detected in the PpASPG1 protein (Figure 5A, time 24, 48, and 72). Figure 5B shows that protein processing was correlated with ASPG activity. At time 0, the activity of the D3 variant was almost negligible whereas for PpASPG1 it was approximately 4 U/mg. The specific activity did not increase with incubation time for PpASPG1, but gradually increased for PpASPG1-D3.

Functional and structural effects of removing the VL were evaluated by comparing substrate preference and thermal stability of PpASPG1 and PpASPG1-D3 (Figure 6). Removal of the VL caused a shift in the profile of substrate preference (Figure 6A) mainly as a result of changes in the K_m and V_{max} values for L-Asn and β -L-Asp-L-Ala (Supplementary Table S3). Consequently, the enzyme efficiency for L-Asn is dramatically reduced. No changes were observed in optimum pH with a value of 8, but the optimum temperature differed from 50°C for PpASPG1 to 35°C for PpASPG1-D3. The activation energy was

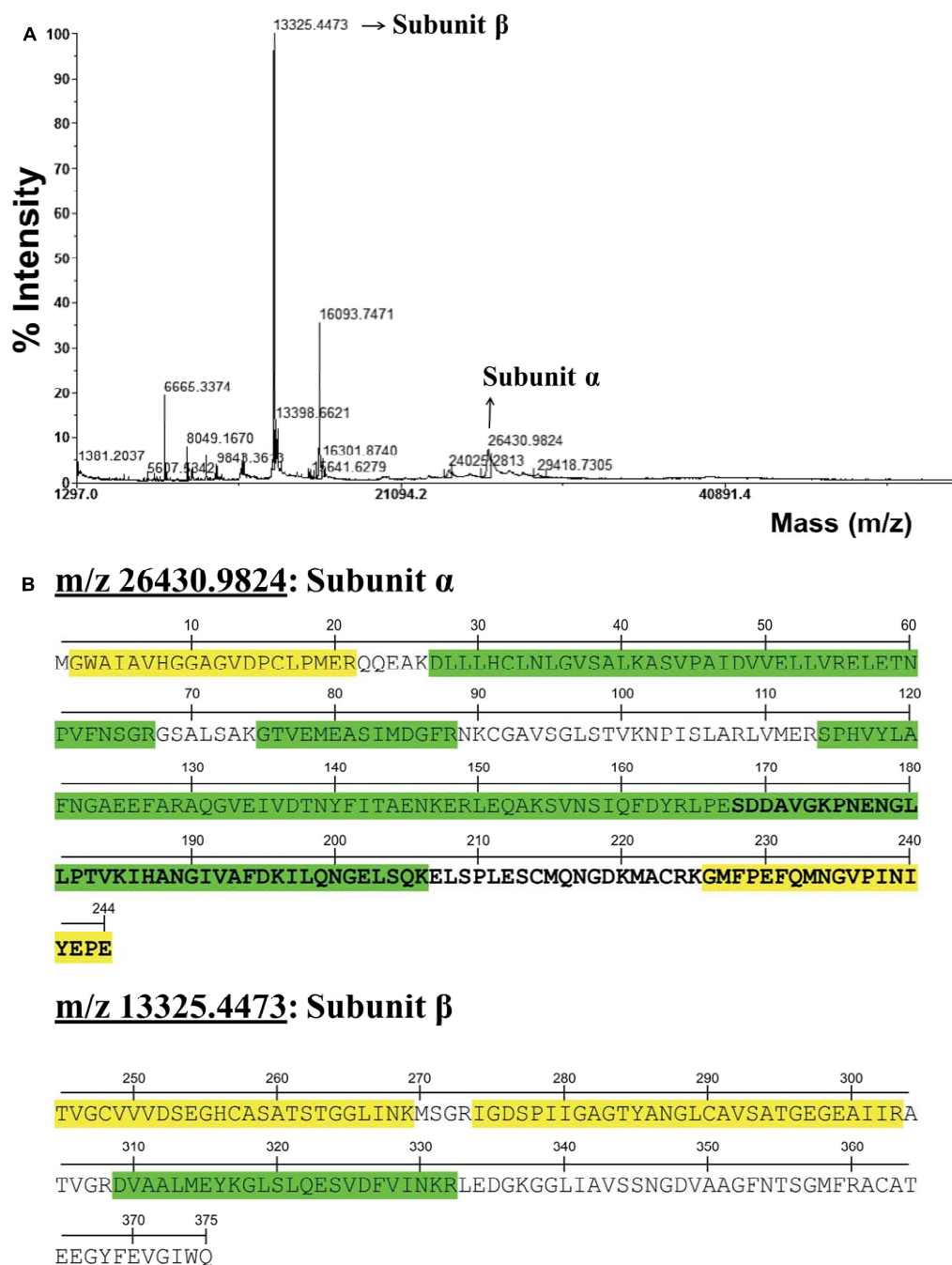


FIGURE 3 | PpASPG1 protein analysis by mass spectrometry. **(A)** Spectrum obtained for undigested purified recombinant PpASPG1 protein by MALDI-TOF analysis in linear mode. **(B)** Identification of the peptides resulting from in gel digestion of the spots corresponding to m/z 26430,9824 and 13325,4473, marked upon the PpASPG1 protein sequence separated by its subunits. Green shaded letters: peptides identified by MALDI-TOF; yellow shaded letters: peptides that were fragmented and their sequences were confirmed by MS/MS; letters in bold: variable sequence region not conserved between homologous ASPG.

27.98 kJ/mol for PpASPG1 and 34.24 kJ/mol for PpASPG1-D3 (Supplementary Figure S2). Furthermore, removal of the VL affected the structural stability of the protein (Figure 6B), since the PpASPG1-D3 variant was clearly more susceptible to thermal inactivation. PpASPG1 activity was not affected after incubation 30 min at 50°C, and retained 75% of initial value after 30 min at

60°C. Conversely, PpASPG1-D3 activity was reduced to half of the starting value after 30 min at 50°C and very low activity was detectable after 30 min at 60°C.

Although, according to the primary structure, PpASPG1 is classified as a K⁺-dependent asparaginase, the substitution of serine for valine in the activation motif prompted us to analyze

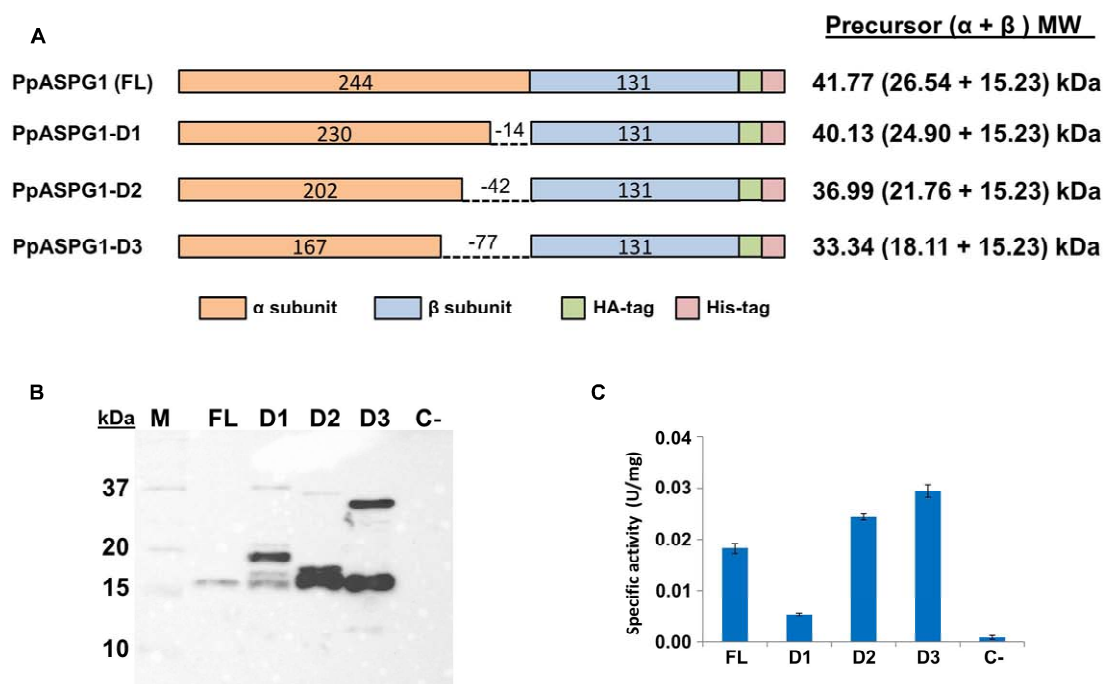


FIGURE 4 | Production of PpASPG1 recombinant proteins in *N. benthamiana* with different deletions in the variable region. **(A)** Schematic representation of the different variants of PpASPG1 that were transiently expressed in tobacco leaves. Variants D1, D2, and D3 were expressed as a precursor from which several amino acids in the variable region were deleted, represented by a dashed line. The number of amino acids is included in the colored bars and the theoretical molecular mass corresponding to the precursor and subunits that would result from correct cleavage is indicated on the right. **(B)** Western blot analysis of the tobacco leaf extracts expressing the different versions of PpASPG1. An antibody against the HA-tag was used and 50 μ g of proteins were loaded in each lane. Protein extract of a leaf that was infiltrated with the empty pJL-TRBO expression vector was also analyzed as a control (C-). **(C)** ASPG activity was determined in the protein extracts. Values are the means of three replicates \pm SD. All values are significantly different according to *t*-tests ($p < 0.01$).

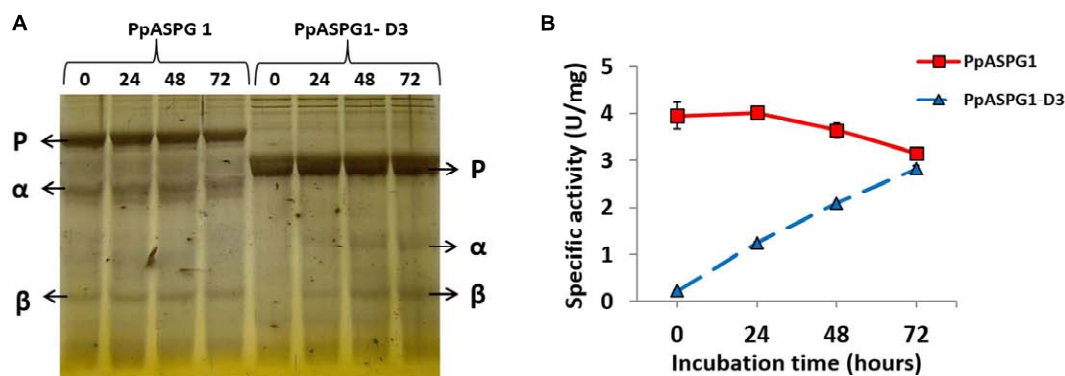


FIGURE 5 | Elimination of the VL allows *in vitro* maturation of the PpASPG1 precursor. Both recombinant proteins were expressed in *E. coli* cells and recovered from inclusion bodies. The purified recombinant proteins were incubated at room temperature during three consecutive days in the presence of KCl 50 mM and aliquots were removed every 24 h and stored at -20°C until analysis. **(A)** Staining of the electrophoretically separated recombinant proteins. Each lane contains 5 μ g of protein and the incubation time in hours is marked on top of the lane. **(B)** Remaining ASPG activity was measured in the 24-h interval aliquots for both proteins. Values are the means of three independent replicates \pm SD. The specific activity values of PpASPG1 at 24, 48, and 72 h are not significantly different of the respective previous value according to *t*-tests ($p < 0.01$). The specific activity values of PpASPG1-D3 at 24, 48, and 72 h are significantly different of the respective previous value according to *t*-tests ($p < 0.01$).

the effect of monovalent cations on the *P. pinaster* enzyme. Asparaginase activities for PpASPG1 and the PpASPG1-D3 variant were compared in the absence or presence of various monovalent cations. As shown in **Figure 7A**, the addition of

Li^+ , Na^+ , and Rb^+ cations did not affect the activity of both enzymes compared to the same reaction lacking any monovalent cation. However, a slight increase in activity was registered when K^+ was added. We also determined the thermal stability of

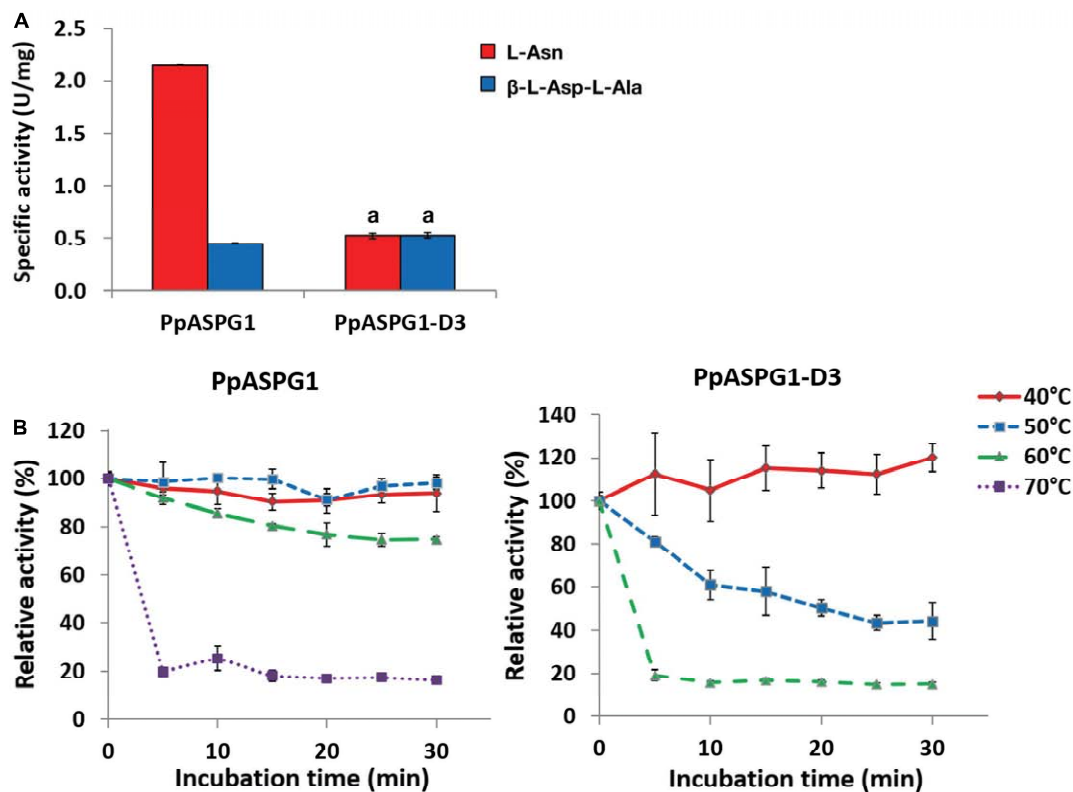


FIGURE 6 | Effect of removing the VL on substrate preference and thermal stability of PpASPG1. Recombinant proteins PpASPG1 and PpASPG1-D3 (see Figure 4A for a schematic representation of the structure) were purified from agroinfiltrated tobacco leaves, and compared for substrate preference and thermal stability. **(A)** Catalytic activity was determined in a buffered medium containing 25 mM of the substrates asparagine or isoaspartyl-alanine dipeptide. Values are means of three replicates \pm SD. Values labeled with the same letter are not significantly different according to *t*-tests ($p < 0.01$). The remainder values are significantly different. **(B)** The purified recombinant proteins were incubated at different temperatures. Aliquots were removed every 5 min and kept on ice until residual ASPG activity was measured. Values are means of three replicates \pm SD. For PpASPG1, the residual activity values after incubating at 70°C are significantly different to those observed after incubating at 40°, 50, and 60°C according to *t*-tests ($p < 0.01$). The residual activity values after incubating at 50 or 60°C are significantly different only after incubating 25 and 30 min. The remainder values at the same incubation time are not significantly different. For PpASPG1-D3, all residual activity values after the same incubation time are significantly different according to *t*-tests ($p < 0.01$), except for values after incubating for 5 min at 40 and 50°C.

both proteins in the absence or presence of potassium ions (Figure 7B). The addition of K^+ increased the thermal stability of both proteins, and the effect was more clearly contrasted in the case of the PpASPG1-D3, which has a lower stability to temperature. After 72 h of incubation at 37°C, no activity was detected for PpASPG1-D3 in the absence of the monovalent ion, whereas 62% of the initial activity was retained in the presence of K^+ at 50 mM.

DISCUSSION

The *P. pinaster* Transcriptome Contains At Least Three ASPG

Most of the ASPG that have been studied in plants belong to angiosperm species and little is known about these enzymes in gymnosperms or primitive terrestrial plants. Genomes of bryophytes include various ASPG genes and their sequences are approximately 57% identical to those of ASPG in vascular plants. However, our phylogenetic analysis (Figure 1) did not assign

these sequences to either K^+ -dependent or K^+ -independent subgroups. In *Selaginella moellendorffii* only one ASPG gene was found which is more similar to the K^+ -independent ASPG. In databases of *Ceratopteris richardii*, *Ginkgo biloba*, and *Cycas rumphii* we did not find sequences with a significant similarity to pine ASPG probably due to low gene expression. Recently, the genomes of several conifer species have been published including *Picea glauca* (Birol et al., 2013; Warren et al., 2015), *P. abies* (Nystedt et al., 2013), and *P. taeda* (Neale et al., 2014; Zimin et al., 2014). Although the genome sequence for *P. pinaster* is not yet available, a reference transcriptome (Canales et al., 2014), contains three complete cDNA sequences for plant-type ASPG which can be classified as K^+ -dependent (PpASPG1) and -independent (PpASPG2 and PpASPG3) according to their similarity with angiosperms sequences. Both *P. abies* and *P. taeda* genomes also contain three ASPG genes with high similarity to those of *P. pinaster*. Unfortunately, with the exception of one gene in *P. taeda* all sequences contained partial ORFs. Therefore, the phylogenetic tree (Figure 1) includes few gymnosperm entries, but the available information suggests that

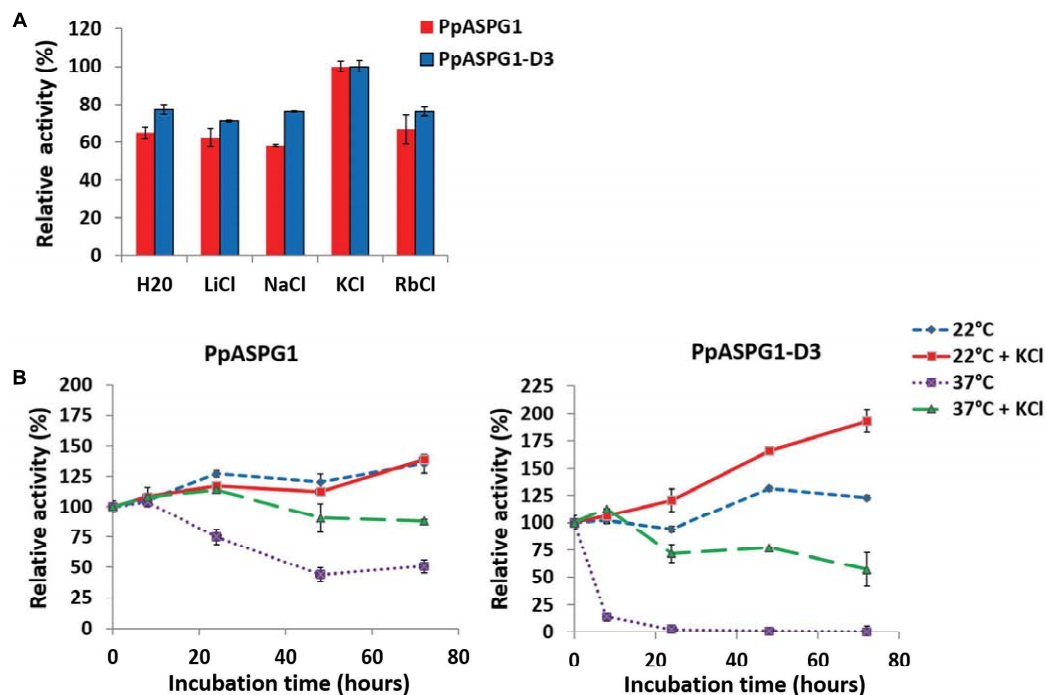


FIGURE 7 | Effect of monovalent cations on enzymatic activity and thermal stability of both PpASPG1 variants. **(A)** ASPG activity was measured for PpASPG1 and PpASPG1-D3 in a buffered medium containing 20 mM L-Asn and 50 mM of the corresponding chloride salt for each ion or water. Values are presented as the mean of three replicates \pm SD. For both PpASPG1 and PpASPG1-D3, the activity values in the presence of KCl are significantly different of those in the different conditions assayed, according to *t*-tests ($p < 0.01$). **(B)** Both protein variants were incubated during three consecutive days at 22 or 37°C in the buffered medium with or without the addition of 50 mM KCl. Aliquots were removed at different time points and residual ASPG activity was assayed. Values are presented as the mean of three replicates \pm SD. For PpASPG1, in the absence of KCl, the activity values were significantly different between 22 and 37°C from the measurements after 24 h of incubation, according to *t*-test ($p < 0.01$). When KCl was present, the activity values are significantly different only at 48 and 72 h. At 22°C, there is not significantly differences in activity values at any incubation time between samples with or without KCl added. At 37°C activity values are significantly different between reactions with or without KCl added at 24, 48, and 72 h. For PpASPG1-D3, activity values are significantly different in the presence or absence of KCl at all incubation times both at 22 and 37°C, except at 8 h for 22°C, according to *t*-tests ($p < 0.01$). Similarly, activity values after incubation at 22 or 37°C are significantly different at all incubation times both in the absence or presence of KCl, except at 8 h with KCl.

conifer genomes contain at least one gene coding for K^+ -dependent ASPG and two for the K^+ -independent type. Several angiosperm genomes have one ASPG gene of each subgroup and the phylogenetic tree also suggests events of genome duplications in some species that contain a higher number of genes. Our analysis based on conservation of primary structure suggests that the separation between subgroups of ASPG is related to spermatophytes, although genome sequences of more Pteridophyta and non-vascular species would be required in the analysis to confirm this hypothesis. No polyploidization events are known for any of the extant pine species (Morse et al., 2009) and whether the presence of two isoforms of the K^+ independent type within conifers is related to a functional specialization remains to be investigated.

All Three Pine ASPG Have Double Activity Asparaginase/Isoaspartyl Dipeptidase

Most ASPG that have been studied in angiosperms belong to the K^+ -independent type (Chang and Farnden, 1981; Lough et al., 1992; Casado et al., 1995; Hejazi et al., 2002; Borek

et al., 2004). As far as we know K^+ -dependent ASPG have been characterized in only four species: *Pisum sativum* (Sodek and Lea, 1993), *A. thaliana* (Bruneau et al., 2006; Gabriel et al., 2012), *L. japonicus* (Credali et al., 2011) and *P. vulgaris* (Bejger et al., 2014). Different findings have been published about the biochemical characteristics of plant ASPG. Characterization of *Lupinus luteus* K^+ -independent ASPG and *E. coli* plant-type EcAIII suggested a main function as isoaspartyl peptidases, while Asn hydrolysis is of secondary importance (Borek et al., 2004). Similar results were reported for the human plant-type asparaginase (hASNase3), which also presents ASPG activity but has more affinity toward isoaspartyl dipeptides (Cantor et al., 2009). The first characterizations of both subtypes within the same species in angiosperms indicated that only the K^+ -independent ASPG has a double catalytic activity while K^+ -dependent ASPG only hydrolyzes Asn (Bruneau et al., 2006; Credali et al., 2011). However, later analysis of *A. thaliana* enzymes showed that the K^+ -dependent ASPG also has hydrolytic activity for isoaspartyl dipeptides (Gabriel et al., 2012), which is in agreement with our observations with the pine enzyme.

Analysis of the recombinant pine ASPG proteins produced in *N. benthamiana* indicates that the three enzymes are able to hydrolyze both substrates, Asn and isoaspartyl dipeptides (Figure 2C), but with differential substrate preference. PpASPG3 shows a clear preference for isoaspartyl alanine with a higher V_{\max} as well as a higher affinity for this substrate (Supplementary Table S3); it catalyzes β -L-Asp-L-Ala hydrolysis almost 22 times more efficiently than Asn hydrolysis. Its characteristics are comparable with those described for K^+ -independent ASPG, indicating that its main function is related to degradation of proteins with β -L-isoaspartyl peptide bonds. In *L. japonicus*, the K^+ -independent ASPG had a K_m of 30 mM for Asn and 0.54 mM for β -Asp-His (Credali et al., 2011) and the orthologous ASPG in *L. luteus* exhibits a K_m value of 4.8 mM for Asn and 0.14 mM for β -Asp-Leu (Borek et al., 2004). For PpASPG2 the V_{\max} values for both substrates were similar and very low so that the collected data could not be fitted to a Michaelis-Menten model. However, we could not rule out that the preferred substrate is an isoaspartyl dipeptide different to β -L-Asp-L-Ala. PpASPG1 was almost 19 times more efficient with Asn than with isoaspartyl alanine, confirming that Asn hydrolysis is the main function of this enzyme, as described for its orthologous enzymes in angiosperms. However, our data also show that PpASPG1 is able to catalyze less efficiently the hydrolysis of isoaspartyl dipeptides, as shown for the product of At3g16150 gene (Gabriel et al., 2012).

The Mature PpASPG1 Enzyme Retains the Complete Variable Region

Plant ASPG, as Ntn-hydrolase family members, are expressed as precursors that undergo a proteolytic activation process placing the catalytic nucleophile residue at the N-terminal of the β subunit. In PpASPG1 and other conifer ASPG, the VL is remarkably long and spans 77 amino acids for which we did not find any sequence similarity in ASPG of non-conifer eukaryotes, except for the last 13 amino acids preceding the nucleophile. Results of peptide fingerprinting and MS/MS show that the complete sequence of the α subunit, including the VL, is present in the mature active enzyme, except for Met1 (Figure 3B). In *L. japonicas*, the VL of ASPG was also present in the mature enzyme (Credali et al., 2011) and it is not unusual for eukaryote expressed proteins to have the N-terminal methionine eliminated (Meinzel et al., 2005). In addition, the m/z analysis indicated a possible sulfoxidation of Met in positions 19, 79, 84, 227, and 233 and oxidation of Trp in position 3. These posttranslational modifications may account for the differences between the m/z of the peptides and their theoretically calculated molecular weight.

When comparing the length of the sequence, some ASPG of bryophytes (PhPaten2 and PhPaten3) also have a long region of approximately 60 residues, but the sequence does not share similarity with that of PpASPG1. The longer VL sequence in the pine enzyme does not seem to have an impact on the mature enzyme, since its function is similar to that of ASPG in angiosperms with a shorter VL. Nevertheless, conservation of VL in conifers may have a

role or beneficial effect related to maturation, structure, or enzyme activity. For this reason, we compared the biochemical characteristics of PpASPG1 and mutant proteins with deletions in the VL.

Effects of VL on Maturation of PpASPG1

The activation of plant ASPG has been addressed by various authors and described as an autoproteolytic maturation process in *A. thaliana* (Bruneau et al., 2006; Gabriel et al., 2012), *L. luteus* (Michalska et al., 2006), *L. japonicus* (Credali et al., 2011), and *P. vulgaris* (Bejger et al., 2014). In contrast, the ASPG from *P. sylvestris* expressed in *E. coli* did not show autocatalytic activity *in vitro* (Cañas et al., 2007). Because of the lack of electron density at the C-terminal of the α subunit of crystallized K^+ -independent ASPG from *L. luteus* (Michalska et al., 2006) and K^+ -dependent ASPG from *P. vulgaris* (Bejger et al., 2014), the role of the VL sequence in the process of autocatalytic activation, substrate binding and catalysis and K^+ activation cannot be evaluated by protein modeling. Therefore, we decided to compare *in planta* the autoproteolysis efficiency of four variants of precursors containing gradually shorter VL. Deletions were performed keeping glutamate as the last α -subunit residue hence the scissile bond to be cleaved for activation in each case was Glu(x)-Thr245. When the different recombinant PpASPG1 variants were expressed in *N. benthamiana* leaves only the unmodified protein was completely processed, although none of the mutations caused total suppression of the activation process. Furthermore, these protein variants produced some aberrantly processed peptides. These results suggest that the full sequence of VL is required for an efficient proteolysis and correct positioning of the activation cleavage.

It is generally observed that, when expressed in *E. coli* cells, the autoproteolysis of plant ASPG is a slow process that continues *in vitro* (Michalska and Jaskólski, 2006; Cantor et al., 2009; Nomme et al., 2012). To evaluate the autoproteolytic capacity of conifer K^+ -dependent ASPG, PpASPG1, and PpASPG1-D3 were expressed in *E. coli*, and both proteins were recovered and purified from inclusion bodies, primarily in their inactive unprocessed precursor form. However, a certain amount of protein processed *in vivo* was observed in the case of PpASPG1, whereas in the case of PpASPG1-D3 this was not detected, which agrees with the higher processing efficiency of PpASPG1 observed when both proteins were produced *in planta*. After 3 days of incubation at room temperature, no signs of *in vitro* cleavage were detected for PpASPG1 whereas the PpASPG1-D3 form did show autoproteolytic activity. Our results indicate that maturation of PpASPG1 might require the assistance of another protein or certain specific conditions as suggested by Cañas et al. (2007) and apparently the VL poses an obstacle for the autocatalytic activation *in vitro*. Because the sequence flanking the nucleophile is not conserved between Ntn-hydrolases, there may be some differences in the particular autoproteolytic process between orthologs (Michalska et al., 2008; Nomme et al., 2012; Su et al., 2013). The VL of angiosperm ASPG is supposed to be unstructured (Bejger et al., 2014), but no information is available for the sequence that is unique to conifer K^+ -dependent ASPG. Structure analysis by crystallization would provide more

insight about the autocatalytic activation of ASPG1 and more experiments are required to determine if the VL could provide a posttranslational control mechanism for enzyme activation in conifers.

The VL Influences Some Biochemical Characteristics of PpASPG1

It has been suggested that the VL might be involved in substrate preference of plant ASPG. In *Arabidopsis*, chimeric proteins with the VL of K⁺-dependent ASPG exchanged for the loop of its K⁺-independent isoform had reduced autoproteolytic capacity and V_{\max} but the K_m value for Asn and isoaspartyl dipeptides was hardly affected. The reciprocal recombinant protein carrying the K⁺-dependent loop in the K⁺-independent backbone had little effect on autoproteolysis but increased the efficiency of Asn hydrolysis by 60-fold (Gabriel et al., 2012). PpASPG1 clearly showed a higher preference for Asn as a substrate than for β -L-Asp-L-Ala by almost 19-fold (Figure 2C and Supplementary Table S3). However, elimination of VL in PpASPG1-D3 determined that this enzyme hydrolyzed Asn and isoaspartyl alanine with almost equal efficiency. In addition, V_{\max} of the modified protein was only about 1/4 of the value for PpASPG1, but this could be because of a large quantity of the purified recombinant protein that remained inactive in its unprocessed precursor form. This suggests that the VL largely reduces the binding capacity of β -L-Asp-L-Ala in the catalytic center in contrast to the smaller Asn molecule, or possibly it hinders the entrance to the active site cavity as was suggested for ASPG from yellow lupin (Michalska et al., 2006).

The substantial differences observed in the optimum temperature and activation energy for PpASPG1 and PpASPG1-D3 suggested that the VL somehow protects the enzyme against thermal degradation and reduces the activation energy. In fact, incubation at different temperatures showed that PpASPG1 enzyme is more stable than PpASPG1-D3 (Figure 6B). The length of these linker sequences has evolved from short in bacteria and Archaea to long in eukaryotes as genome complexity increases, although plants generally have the shortest linker sequences within eukaryote organisms (Wang et al., 2011). Possibly the long VL of PpASPG1 could imply advantages in the stability and folding of the enzyme.

Enhancement of Activity by Potassium Cations in PpASPG1 Is Limited

PpASPG1 has a higher similarity to the K⁺-dependent type ASPG in angiosperms (Figure 1). Both, PpASPG1 and PpASPG1-D3 were most active when K⁺ was included in the buffer (Figure 7A), which would validate the classification of PpASPG1 within the K⁺-dependent group of ASPG. However, in the presence of other ions or water they maintained 65% to 75% of their catalytic activity, which makes them far less dependent of K⁺ than their orthologs in angiosperms which suffer a 90% loss of activity in absence of the cation (Bruneau et al., 2006; Credali et al., 2011). The three amino acids that were referred to as the catalytic switch (His117, Arg224, and Glu250 in the PvAspG1 protein) by Bejger

et al. (2014) are conserved amongst all sequences included in the alignment of Supplementary Figure S1 and also among all the sequences included in the phylogenetic tree (Figure 1). Nevertheless, it is the binding of the metal ion in the activation loop what induces the conformational changes in these residues setting this catalytic switch to ON or OFF to allow or prevent anchoring of the reaction substrate/product in the active site (Bejger et al., 2014). Therefore, the differential effect of K⁺ ion on the conifer enzyme may be as a result of differences in the sequence of the activation loop. In the PvAspG1 protein the activation loop is formed by eight residues (VMDKSPHS) (see Supplementary Figure S1). In particular, Ser118 seems to be of key importance for the potassium dependence of PvAspG1 and it is substituted in PpASPG1 and in the independent K⁺-enzymes by a hydrophobic amino acid (valine or isoleucine). This residue is responsible for the proper positioning of water molecules coordinated by the metal ion and also mediates the conformational changes of Glu250, one of the catalytic switch residues, effected by the exchange of the coordinated alkali-metal cation (Bejger et al., 2014). In fact, this serine residue is absolutely conserved in all K⁺-dependent proteins in the phylogenetic tree on Figure 1 with the exception of the conifer protein. Consequently, this difference in the activation loop of the pine enzyme may prevent the proper functioning of the ON/OFF K⁺-dependent mechanism described for its orthologs in angiosperms.

Since PpASPG1-D3 only differs from the PpASPG1 complete sequence in lacking the VL and the activation loop is kept intact, the effect of the metal ion was expected to be similar for both proteins. We found a larger difference in relative activity with/without K⁺ for PpASPG1 than for PpASPG1-D3, but this could be due to the higher stability of the enzyme carrying the VL. Incubation of the two recombinant ASPG during 3 days in the presence and absence of K⁺ suggests that the metal ion contributes in maintaining the functional integrity of the enzyme (Figure 7B) as also reported for pea ASPG (Sodek, 1980). In fact, this protective effect against denaturalization might explain the higher activity of the PpASPG1 in the presence of K⁺ instead of an actual activation mechanism as described in angiosperms.

AUTHOR CONTRIBUTIONS

FMC and FRC designed the experiments, FdT, RC, and CA contributed in the design of some experiments. SVK and FdT performed the experiments. SVK, FMC, and FRC analyzed the data and wrote the manuscript. All authors have read and approved the final manuscript.

FUNDING

This work was supported by grants from the Spanish Ministerio de Economía y Competitividad (BIO2015-69285-R), Ministerio de Educación (AP2010-5434) and Junta de Andalucía (BIO-474). RAC was supported by the grant BIO2015-73512-JIN (MINECO/AEI/FEDER, UE).

ACKNOWLEDGMENT

We are grateful to Dr. Sophien Kamoun from The Sainsbury Laboratory for providing us with the pJL-TRBO vector.

REFERENCES

- Azevedo, R. A., Lancien, M., and Lea, P. J. (2006). The aspartic acid metabolic pathway, an exciting and essential pathway in plants. *Amino Acids* 30, 143–162. doi: 10.1007/s00726-005-0245-2
- Bejger, M., Imiolczyk, B., Clavel, D., Gilski, M., Pajak, A., Marsolais, F., et al. (2014). Na⁺/K⁺ exchange switches the catalytic apparatus of potassium-dependent plant L-asparaginase. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 70, 1854–1872. doi: 10.1107/S1399004714008700
- Biol, I., Raymond, A., Jackman, S. D., Pleasance, S., Coope, R., Taylor, G. A., et al. (2013). Assembling the 20 Gb white spruce (*Picea glauca*) genome from whole-genome shotgun sequencing data. *Bioinformatics* 29, 1492–1497. doi: 10.1093/bioinformatics/btt178
- Borek, D., and Jaskólski, M. (2001). Sequence analysis of enzymes with asparaginase activity. *Acta Biochim. Pol.* 48, 893–902.
- Borek, D., Michalska, K., Brzezinski, K., Kisiel, A., Podkowinski, J., Bonthron, D. T., et al. (2004). Expression, purification and catalytic activity of *Lupinus luteus* asparagine β -amidohydrolase and its *Escherichia coli* homolog. *Eur. J. Biochem.* 271, 3215–3226. doi: 10.1111/j.1432-1033.2004.04254.x
- Bradford, M. (1976). A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3
- Brannigan, J., Dodson, G., Duggleby, H., Moody, P., Smith, J., Tomchick, D., et al. (1995). A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature* 378, 416–419. doi: 10.1038/378416a0
- Bruneau, L., Chapman, R., and Marsolais, F. (2006). Co-occurrence of both L-asparaginase subtypes in Arabidopsis: At3g16150 encodes a K⁺-dependent L-asparaginase. *Planta* 224, 668–679. doi: 10.1007/s00425-006-0245-9
- Canales, J., Ávila, C., Cantón, F. R., Pacheco-Villalobos, D., Díaz-Moreno, S., Ariza, D., et al. (2012). Gene expression profiling in the stem of young maritime pine trees: detection of ammonium stress-responsive genes in the apex. *Trees Struct. Funct.* 26, 609–619. doi: 10.1007/s00468-011-0625-z
- Canales, J., Bautista, R., Label, P., Gómez-Maldonado, J., Lesur, I., Fernández-Pozo, N., et al. (2014). De novo assembly of maritime pine transcriptome: implications for forest breeding and biotechnology. *Plant Biotechnol. J.* 12, 286–299. doi: 10.1111/pbi.12136
- Cañas, R. A., de la Torre, F., Cánovas, F. M., and Cantón, F. R. (2007). Coordination of PsAS1 and PsASPG expression controls timing of re-allocated N utilization in hypocotyls of pine seedlings. *Planta* 225, 1205–1219. doi: 10.1007/s00425-006-0431-9
- Cantor, J. R., Stone, E. M., Chantranupong, L., and Georgiou, G. (2009). The human asparaginase-like protein 1 hASRGL1 is an Ntn hydrolase with β -aspartyl peptidase activity. *Biochemistry* 48, 11026–11031. doi: 10.1021/bi901397h
- Casado, A., Caballero, J. L., Franco, A. R., Cárdenas, J., Grant, M. R., and Muñoz-Blanco, J. (1995). Molecular cloning of the gene encoding the L-asparaginase gene of *Arabidopsis thaliana*. *Plant Physiol.* 108, 1321–1322. doi: 10.1104/pp.108.3.1321
- Chang, K. S., and Farnden, K. J. F. (1981). Purification and properties of asparaginase from *Lupinus arboreus* and *Lupinus angustifolius*. *Arch. Biochem. Biophys.* 208, 49–58. doi: 10.1016/0003-9861(81)90122-3
- Credali, A., Díaz-Quintana, A., García-Calderón, M., De la Rosa, M. A., Márquez, A. J., and Vega, J. M. (2011). Structural analysis of K⁺ dependence in L-asparaginases from *Lotus japonicus*. *Planta* 234, 109–122. doi: 10.1007/s00425-011-1393-0
- Gabriel, M., Telmer, P. G., and Marsolais, F. (2012). Role of asparaginase variable loop at the carboxyl terminal of the alpha subunit in the determination of substrate preference in plants. *Planta* 235, 1013–1022. doi: 10.1007/s00425-011-1557-y
- Galperin, M. Y., and Koonin, E. V. (2012). Divergence and convergence in enzyme evolution. *J. Biol. Chem.* 287, 21–28. doi: 10.1074/jbc.R111.241976

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2017.01075/full#supplementary-material>

- Gaufichon, L., Masclaux-Daubresse, C., Tcherkez, G., Reisdorf-Cren, M., Sakakibara, Y., Hase, T., et al. (2013). *Arabidopsis thaliana* ASN2 encoding asparagine synthetase is involved in the control of nitrogen assimilation and export during vegetative growth. *Plant Cell Environ.* 36, 328–342. doi: 10.1111/j.1365-3040.2012.02576.x
- Guan, C., Cui, T., Rao, V., Liao, W., Benner, J., Lin, C. L., et al. (1996). Activation of glycosylasparaginase. Formation of active accession numbers N-terminal threonine by intramolecular autoproteolysis. *J. Biol. Chem.* 271, 1732–1737. doi: 10.1074/jbc.271.3.1732
- Hejazi, M., Piotukh, K., Mattow, J., Deutzmann, R., Volkmer-Engert, R., and Lockau, W. (2002). Isoaspartyl dipeptidase activity of plant-type asparaginases. *Biochem. J.* 364, 129–136. doi: 10.1042/bj3640129
- Hildebrandt, T. M., Nesi, A. N., Araújo, W. L., and Braun, H. (2015). Amino acid catabolism in plants. *Mol. Plant* 8, 1563–1579. doi: 10.1016/j.molp.2015.09.005
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0
- Lea, P. J., Sodek, L., Parry, M. A. J., Shewry, P. R., and Halford, N. G. (2007). Asparagine in plants. *Ann. Appl. Biol.* 150, 1–26. doi: 10.1111/j.1744-7348.2006.00104.x
- Lev-Yadun, S., and Sederoff, R. (2000). Pines as model gymnosperms to study evolution, wood formation, and perennial growth. *J. Plant Growth Regul.* 19, 290–305. doi: 10.1007/s003440000045
- Lindbo, J. A. (2007). TRBO: a high-efficiency *Tobacco mosaic virus* RNA-based overexpression vector. *Plant Physiol.* 145, 1232–1240. doi: 10.1104/pp.107.106377
- Loan, D. K., Con, T. H., Hong, T. T., and Mai Ly, L. T. (2013). Quick determination of ammonia ions in water environment based on thymol color creating reaction. *Environ. Sci.* 1, 83–92. doi: 10.12988/es.2013.31010
- Lough, T. J., Chang, K. S., Carne, A., Monk, B. C., Reynolds, P. H. S., and Farnden, K. J. F. (1992). L-asparaginase from developing seeds of *Lupinus arboreus*. *Phytochemistry* 31, 1519–1527. doi: 10.1016/0031-9422(92)83098-J
- Meinell, T., Peynot, P., and Giglione, C. (2005). Processed N-termini of mature proteins in higher eukaryotes and their major contribution to dynamic proteomics. *Biochimie* 87, 701–712. doi: 10.1016/j.biochi.2005.03.011
- Michalska, K., Bujacz, G., and Jaskólski, M. (2006). Crystal structure of plant asparaginase. *J. Mol. Biol.* 360, 105–116. doi: 10.1016/j.jmb.2006.04.066
- Michalska, K., Hernandez-Santoyo, A., and Jaskólski, M. (2008). The mechanism of autocatalytic activation of plant-type L-asparaginases. *J. Biol. Chem.* 283, 13388–13397. doi: 10.1074/jbc.M800746200
- Michalska, K., and Jaskólski, M. (2006). Structural aspects of L-asparaginases, their friends and relations. *Acta Biochim. Pol.* 53, 627–640.
- Möller, H. (1985). “L-aspartate and L-asparagine,” in *Methods of Enzymatic Analysis*, 3rd Edn, Vol. VIII, eds H. U. Bergmeyer, J. Bergmeyer, and I. M. Grass (Weinheim: VCH Verlagsgesellschaft), 350–357.
- Morse, A. M., Peterson, D. G., Islam-faridi, M. N., Smith, K. E., Magbanua, Z., Garcia, S. A., et al. (2009). Evolution of genome size and complexity in Pinus. *PLoS ONE* 4:e4332. doi: 10.1371/journal.pone.0004332
- Neale, D. B., Wegrzyn, J. L., Stevens, K. A., Zimin, A. V., Puiu, D., Crepeau, M. W., et al. (2014). Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biol.* 15:R59. doi: 10.1186/gb-2014-15-3-r59
- Nome, J., Su, Y., Konrad, M., and Lavie, A. (2012). Structures of apo and product-bound human L-asparaginase: insights into the mechanism of autoproteolysis and substrate hydrolysis. *Biochemistry* 51, 6816–6826.
- Nystedt, B., Street, N. R., Wetterborn, A., Zuccolo, A., Lin, Y. C., Scofield, D. G., et al. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature* 497, 579–584. doi: 10.1038/nature12211
- Salvador, L., Alía, R., Agúndez, D., and Gil, L. (2000). Genetic variation and migration pathways of maritime pine (*Pinus pinaster* Ait) in the Iberian peninsula. *TAG Theor. Appl. Genet.* 100, 89–95. doi: 10.1007/s001220050013

- Sieciechowicz, K. A., Joy, K. W., and Ireland, R. J. (1988). The metabolism of asparagine in plants. *Phytochemistry* 27, 663–671. doi: 10.1016/0031-9422(88)84071-8
- Sodek, L. (1980). Distribution and properties of a potassium-dependent asparaginase isolated from developing seeds of *Pisum sativum* and other plants. *Plant Physiol.* 65, 22–26. doi: 10.1104/pp.65.1.22
- Sodek, L., and Lea, P. J. (1993). Asparaginase from the testa of developing lupin and pea seeds. *Phytochemistry* 34, 51–56. doi: 10.1016/S0031-9422(00)90781-7
- Su, Y., Karamitros, C. S., Nomme, J., McSorley, T., Konrad, M., and Lavie, A. (2013). Free glycine accelerates the autolytic activation of human asparaginase. *Chem. Biol.* 20, 533–540. doi: 10.1016/j.chembiol.2013.03.006
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Villalobos, D. P., Díaz-Moreno, S. M., Said, E.-S. S., Cañas, R. A., Osuna, D., Van Kerckhoven, S. H. E., et al. (2012). Reprogramming of gene expression during compression wood formation in pine: coordinated modulation of S-adenosylmethionine, lignin and lignan related genes. *BMC Plant Biol.* 12:100. doi: 10.1186/1471-2229-12-100
- Wang, M., Kurland, C. G., and Caetano-Anollés, G. (2011). Reductive evolution of proteomes and protein structures. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11954–11958. doi: 10.1073/pnas.1017361108
- Warren, R. L., Keeling, C. I., Yuen, M. M., Raymond, A., Taylor, G. A., Vandervalk, B. P., et al. (2015). Improved white spruce (*Picea glauca*) genome assemblies and annotation of large gene families of conifer terpenoid and phenolic defense metabolism. *Plant J.* 83, 189–212. doi: 10.1111/tj.12886
- Xu, Q., Buckley, D., Guan, C., and Guo, H. C. (1999). Structural insights into the mechanism of intramolecular proteolysis. *Cell* 98, 651–661. doi: 10.1016/S0092-8674(00)80052-5
- Zimin, A., Stevens, K. A., Crepeau, M. W., Holtz-Morris, A., Koriabine, M., Marçais, G., et al. (2014). Sequencing and assembly of the 22-Gb loblolly pine genome. *Genetics* 196, 875–890. doi: 10.1534/genetics.113.159715

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Van Kerckhoven, de la Torre, Cañas, Avila, Cantón and Cánovas. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

RESEARCH ARTICLE

Open Access

Reprogramming of gene expression during compression wood formation in pine: Coordinated modulation of S-adenosylmethionine, lignin and lignan related genes

David P Villalobos^{1,2}, Sara M Díaz-Moreno^{1,3}, El-Sayed S Said¹, Rafael A Cañas¹, Daniel Osuna^{1,4}, Sonia H E Van Kerckhoven¹, Rocío Bautista¹, Manuel Gonzalo Claros¹, Francisco M Cánovas¹ and Francisco R Cantón^{1*}

Abstract

Background: Transcript profiling of differentiating secondary xylem has allowed us to draw a general picture of the genes involved in wood formation. However, our knowledge is still limited about the regulatory mechanisms that coordinate and modulate the different pathways providing substrates during xylogenesis. The development of compression wood in conifers constitutes an exceptional model for these studies. Although differential expression of a few genes in differentiating compression wood compared to normal or opposite wood has been reported, the broad range of features that distinguish this reaction wood suggest that the expression of a larger set of genes would be modified.

Results: By combining the construction of different cDNA libraries with microarray analyses we have identified a total of 496 genes in maritime pine (*Pinus pinaster*, Ait.) that change in expression during differentiation of compression wood (331 up-regulated and 165 down-regulated compared to opposite wood). Samples from different provenances collected in different years and geographic locations were integrated into the analyses to mitigate the effects of multiple sources of variability. This strategy allowed us to define a group of genes that are consistently associated with compression wood formation. Correlating with the deposition of a thicker secondary cell wall that characterizes compression wood development, the expression of a number of genes involved in synthesis of cellulose, hemicellulose, lignin and lignans was up-regulated. Further analysis of a set of these genes involved in S-adenosylmethionine metabolism, ammonium recycling, and lignin and lignans biosynthesis showed changes in expression levels in parallel to the levels of lignin accumulation in cells undergoing xylogenesis *in vivo* and *in vitro*.

Conclusions: The comparative transcriptomic analysis reported here have revealed a broad spectrum of coordinated transcriptional modulation of genes involved in biosynthesis of different cell wall polymers associated with within-tree variations in pine wood structure and composition. In particular, we demonstrate the coordinated modulation at transcriptional level of a gene set involved in S-adenosylmethionine synthesis and ammonium assimilation with increased demand for coniferyl alcohol for lignin and lignan synthesis, enabling a better understanding of the metabolic requirements in cells undergoing lignification.

* Correspondence: frcanton@uma.es

¹Departamento de Biología Molecular y Bioquímica, Facultad de Ciencias, Universidad de Málaga, Campus Universitario de Teatinos, 29071, Málaga, Spain

Full list of author information is available at the end of the article

Background

Large amounts of wood can be formed throughout the life of a tree through a complex process of cell differentiation called xylogenesis. In this process, cambium-derived cells undergo cell division followed by thickening of the secondary cell wall by modification of the synthesis and deposition of cellulose, hemicelluloses, cell wall proteins and lignin, and finally programmed cell death to develop tracheary elements [1]. Genes involved in these cellular processes are under strict transcriptional regulation during different stages of differentiation [2]. Nevertheless, inputs from external cues are also integrated in this developmental program to adapt secondary xylem properties to growth requirements in a continuously changing environment [3]. As a result of this interaction, natural variations are found in wood properties not only among different species and genotypes, but also within the same tree [4].

Transcriptome analysis of differentiating secondary xylem has allowed us to draw a general picture of the genes, metabolic pathways and potential regulators involved in wood formation [5-11]. However, our knowledge is limited as to how the transcriptomes, proteomes and metabolomes involved in wood development are modulated by developmental and environmental signals to cause within-tree variation in wood properties. Considerable effort has been focused on studying the main pathways that lead to monolignol biosynthesis [12] and carbohydrate partitioning to cellulose [13], as well as understanding how changes in gene expression in those pathways may affect cellular wall characteristics and, consequently, wood quality [14]. However, less attention has been paid to the pathways that provide S-adenosylmethionine (SAM) for methylation reactions during biosynthesis of coniferyl and sinapyl alcohols, even though transcripts and proteins for enzymes of these pathways are abundant in the developing xylem [15,16]. Consumption of methyl units during lignification also implies the existence of an important carbon sink, and modifications in the availability of SAM may affect wood quality through alterations in lignin content and composition [16-18]. Furthermore, it has been proposed that glycine decarboxylase activity associated with SAM metabolism could be an important source of released ammonium, which must be efficiently re-assimilated to prevent significant imbalances in the strict nitrogen economy of the plant [19].

Molecular determinants of intra-genotype variation can be identified by comparing transcriptomes from cambial derivatives undergoing differentiation in the same tree that produce wood that differs in composition and structure [4]. However, the plasticity of the molecular machinery involved in wood formation [3] may result in some difficulties in terms of defining candidate genes when expression patterns are compared over different annual growth periods. Therefore, the definition of a consistent set of

candidate genes for wood property variations requires analysis that mitigates sources of variation from genotypes and local environmental changes.

Reaction wood illustrates how the integration of environmental signals into the secondary xylem developmental program leads to intra-genotype variation in wood. When the stem of a woody plant grows in a non-vertical orientation it forms reaction wood, a specialized secondary xylem that helps the stem maintain a certain orientation and re-orient growth [20]. In gymnosperms, reaction wood is called compression wood and it develops on the underside of branches and inclined stems [21]. Cell walls in compression wood tracheids are thicker than in normal wood and lack the innermost S3 layer. It also contains more lignin, less cellulose and altered levels of hemicelluloses than normal wood. Moreover, in compression wood cellulose microfibrils are deposited with increased angle relative to the cell axis [22,23], and the lignin is enriched in p-hydroxyphenylpropane units [24,25]. In contrast, opposite wood develops on the opposite side of the inclined stem and it is more similar to normal wood [25]. Differential expression of a few genes during compression wood differentiation compared to normal or opposite wood has been reported [16,26-28]. However, the clear anatomical, structural and compositional differences that characterize compression wood suggest that the expression of more genes would be modified during its differentiation [25,29,30].

We are interested in how biosynthetic pathways that provide substrates for xylogenesis are coordinated and regulated according to the different demands during development of wood with distinct composition. In particular, comparison of samples from differentiating compression wood and opposite wood from the same tree allows us to analyze the transcriptome changes accompanying to different levels of lignin deposition during xylogenesis. In this work, we present the results of a comparative transcriptomic analysis that comprehensively uses a range of cDNA libraries and microarray analyses, combining samples from different *Pinus pinaster* provenances collected in different years and geographic locations, to identify changes in the transcriptome associated with compression wood development.

Results

Identification of genes differentially expressed during compression and opposite wood formation in *Pinus pinaster*

To identify genes that are differentially expressed between compression (Cx) and opposite (Ox) developing xylem we followed the strategy depicted in Figure 1. Samples from different provenances collected in different years at two latitudes were integrated into the analyses to mitigate the effects of multiple sources of variability. Two microarrays were manufactured with cDNA libraries constructed with RNA isolated from adult maritime pine trees of Corsican

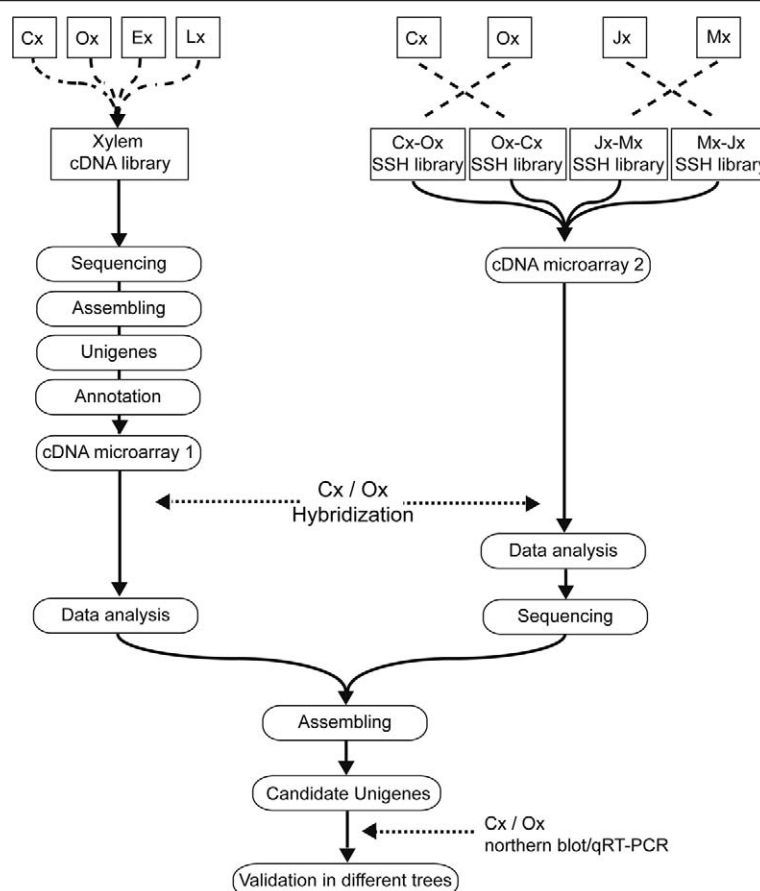


Figure 1 Schematic overview of the experimental procedure used in this work to identify genes differentially expressed between compression and opposite developing xylem from maritime pine. Samples of differentiating secondary xylem forming compression (Cx), opposite (Ox), early (Ex) and late (Lx) developing xylem collected in 1998, 1999 and 2000. The second microarray (Figure 1, cDNA microarray 2) was constructed with 4041 cDNA clones from four subtractive libraries obtained by reciprocally subtracting cDNA populations obtained from Cx and Ox samples that were collected in 1998 and from juvenile (Jx) and mature (Mx) developing xylem collected in 2001. Both microarrays were hybridized with labeled aRNA targets synthesized from Cx and Ox collected from four maritime pine trees of the Sierra Bermeja provenance, Spain [31], sampled in 2005. To validate the expression patterns of selected genes, northern blot and qRT-PCR analyses were carried out with samples of adult trees from Sierra Bermeja provenance collected in 2005 and 2008, respectively.

and Aquitaine provenances grown in Aquitaine, France. The first microarray (Figure 1, cDNA microarray 1) was constructed using a set of 2800 cDNAs from samples of Cx, Ox, early (Ex) and late (Lx) developing xylem collected in 1998, 1999 and 2000. The second microarray (Figure 1, cDNA microarray 2) was constructed with 4041 cDNA clones from four subtractive libraries obtained by reciprocally subtracting cDNA populations obtained from Cx and Ox samples that were collected in 1998 and from juvenile (Jx) and mature (Mx) developing xylem collected in 2001. Both microarrays were hybridized with labeled aRNA targets synthesized from Cx and Ox collected from four maritime pine trees of the Sierra Bermeja provenance, Spain [31], sampled in 2005. To validate the expression patterns of selected genes, northern blot and qRT-PCR analyses were carried out with samples of adult trees from Sierra Bermeja provenance collected in 2005 and 2008, respectively.

Microscopic examination and lignin quantification were performed for phenotypic validation of the samples

(Figure 2). Consistent with the typical morphology of compression wood tracheids [25], transversal sections of developing tracheids showed a rounded shape and many intercellular spaces, whereas Ox developing tracheids had straight sides and few intercellular spaces (Figure 2a). Lignin content was determined in samples from the four trees used in microarray analyses and the four trees used in qRT-PCR (Figure 2b). Cx samples contained significantly more lignin than Ox samples. However, variability in lignin content between both types of developing xylem was evident among trees. In particular, the difference in lignin content between Cx and Ox samples from tree T4-08 was lower compared to the other trees and not statistically significant according to a *t*-test.

Results of hybridization analyses of both microarrays are summarized as volcano plots in Figure 3. After eliminating spots flagged as bad, not found or absent and those with a signal intensity that did not surpass twofold their background signal, only 2598 and 3148 spots were kept for further analysis in microarray 1 and 2 respectively. Relative

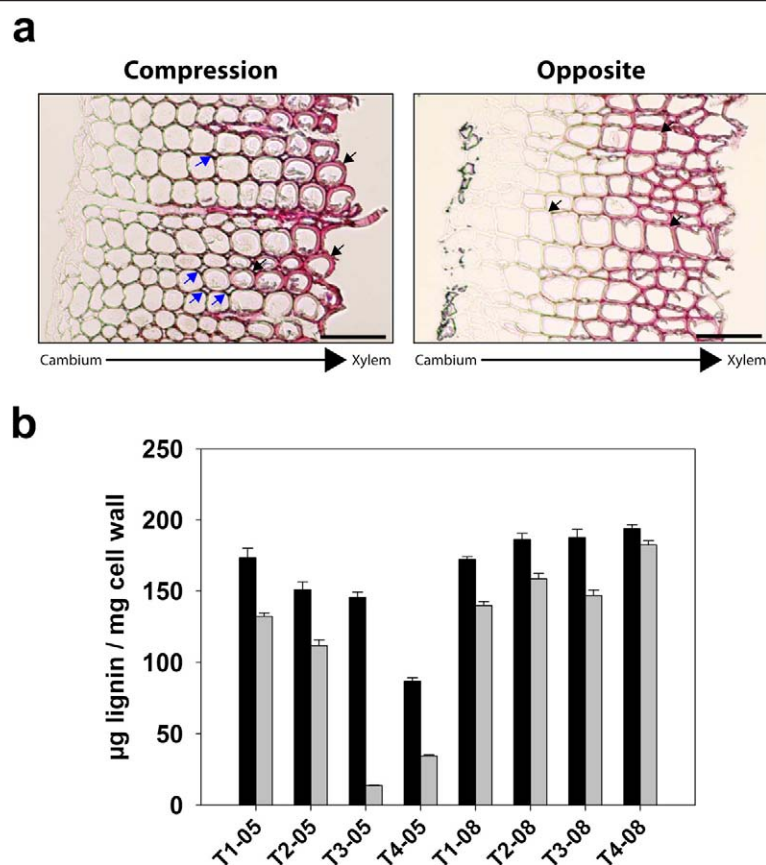


Figure 2 Cross-sectional images and lignin content analysis of cambial scraping samples used in gene expression analyses. (a) Representative microscopic images of thin transversal cross-sections of cambial scrapings stained with phloroglucinol. The orientation of the differentiating tracheids in the scraping relative to the cambium and developed xylem is indicated at the bottom of the pictures. Blue arrows mark typical intercellular spaces in compression wood, and black arrows mark cells with characteristic rounded shape in compression wood or straight sides in opposite wood. Scale bar at the bottom left represents 0.1 mm. (b) Lignin content of cambial scrapings from compression (black bars) and opposite (grey bars) sides of the eight different maritime pine trees analyzed, sampled in May 2005 (T1-05, T2-05, T3-05, T4-05) and May 2008 (T1-08, T2-08, T3-08, T4-08) after two months of artificial mechanical bending. Lines on the top of every bar show standard error of the mean for four independent replicates. Statistical significance of differences in lignin content between cambial scrapings from compression and opposite sides from the same tree was determined with a paired Student's *t*-test ($p < 0.05$). All differences were statistically significant except for tree T4-08.

expression values (\log_2 Fold Change), *p*-values and adjusted *p*-values for all the spots represented in Figure 3 are available in Additional files 1 and 2. A clear bias is observed in both cases, with a higher number of spots showing up-regulation in Cx compared to Ox. The analysis of microarray 1 shows 229 spots containing probes for genes that were significantly up-regulated ($p \leq 0.001$) by at least 1.5-fold during compression wood formation, and only 90 spots showing up-regulation during opposite wood formation (Figure 3a). Since these cDNAs were previously sequenced to construct an EST database of *Pinus pinaster* developing xylem the sequences were already available [32]. The analysis of microarray 2 identified 291 spots showing up-regulation in Cx and 132 spots showing up-regulation in Ox (Figure 3b). Only these 423 cDNA clones were sequenced from the probe set of microarray 2.

Sequences with an average Phred score >20 per nucleotide in a sliding window of 15 nucleotides and a minimal length of 100 nucleotides were kept for further analysis (average length of 437 nucleotides and a standard deviation of ± 171). Those with lower quality scores or shorter than 100 nucleotides were re-sequenced. The complete set of sequences corresponding to differentially expressed genes from microarray 1 and 2 were used to define a set of unigenes, and a total number of 355 unigenes up-regulated in Cx and 176 unigenes up-regulated in Ox were obtained. Moderate values of sequence redundancy were obtained (31.6% for Cx and 21.2% for Ox).

To validate the transcriptomic analysis, a group of genes was selected and their cDNAs were used as probes for northern blot analysis of RNA extracted from Cx

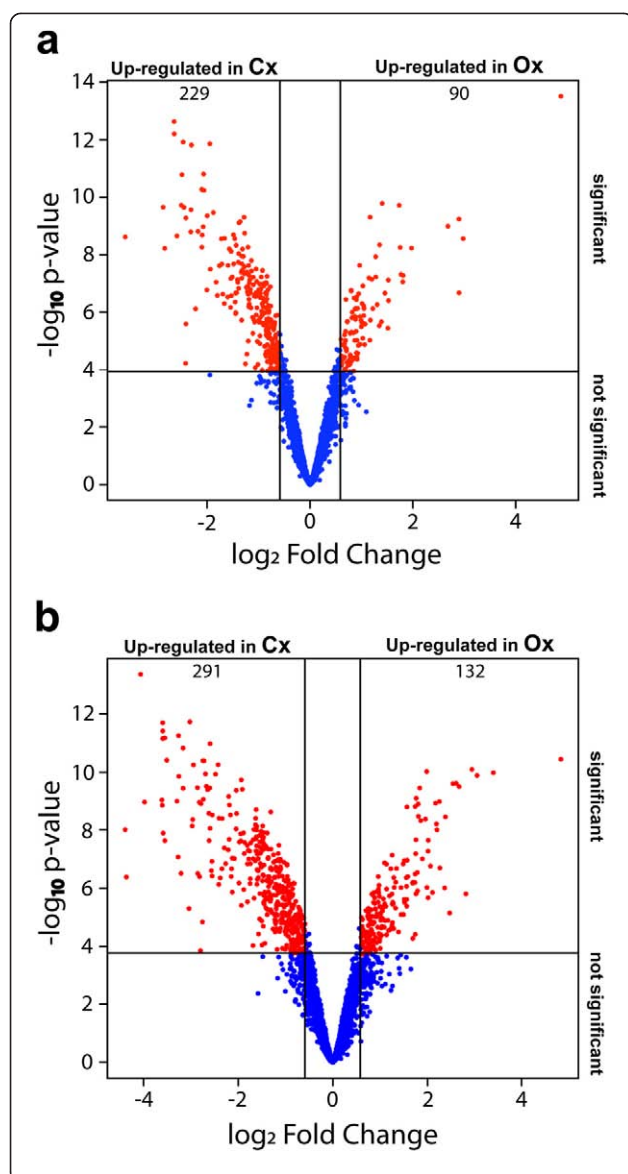


Figure 3 Volcano plots of microarray analyses to identify genes differentially expressed during compression and opposite wood formation. The common logarithm of the p -value was represented as a function of the binary logarithm of the background-corrected and normalized opposite:compression fluorescence ratio (\log_2 Fold Change) for each spot. Vertical bars delimit the spots showing up-regulation in developing compression xylem by at least 1.5-fold compared to developing opposite xylem (Up-regulated in Cx) or spots showing up-regulation in developing opposite xylem by at least 1.5-fold compared to developing compression xylem (Up-regulated in Ox). The horizontal line delimits the spots showing significant up-regulation under the criteria of an adjusted p -value ≤ 0.001 . Therefore, the upper left and right sectors delimited by the horizontal and vertical lines include the spots (in red) containing probes for genes significantly up-regulated in developing compression or opposite xylem respectively. The number of spots corresponding with genes significantly up-regulated in Cx or Ox are shown in the top side of the sector. (a) Results from the analysis of microarray 1 constructed with cDNA clones from the composite library. (b) Results from the analysis of microarray 2 constructed with cDNA clones from subtractive libraries.

and Ox (Figure 4). The differential expression patterns in Cx and Ox observed in microarrays analyses were validated for all selected genes. In particular, northern blot analysis suggested that our strategy allowed us to identify some genes that may be specifically induced at high levels in Cx whereas are not expressed or expressed at very low levels in Ox (Figure 4, *EFE*, *GST*, *GLP*, *XGT*, *LHT*). As a control, the expression pattern was also confirmed for a gene that is expressed at similar levels in both types of differentiating xylem according to microarray analysis (Figure 4b, panel labeled *13-7XLA6*).

Transcriptome changes in functional categories related to cell wall during compression wood formation in pine

The complete set of unigenes was functionally annotated using BlastX analysis [33] against GenBank and BlastN

using the Pine Gene Index database (Additional file 3). Sequences that matched with the same entry in the database were assumed to represent the same gene. Therefore, the final numbers of unigenes were reduced to 331 for Cx and 165 for Ox. Most of these genes showed significant similarities to sequences in databases (293 in Cx and 145 in Ox), although some of them were similar to sequences with unknown function (49 in Cx and 45 in Ox). The number of unigenes with no significant similarity was low in both cases (38 in Cx and 20 in Ox).

The genes with assigned function were grouped into functional categories using the *Arabidopsis thaliana* Munich Information Centre for Protein Sequences (MIPS) database, and suppression of redundancy in MIPS functional assignments by decision according to their most probable role in xylem development (Additional file 3). In keeping with the greater number of genes identified as up-regulated in Cx, most of the functional categories included more genes in this tissue (Figure 5 and Additional file 3 for details). The most represented categories were “C-compound and carbohydrate metabolism” (30 in Cx/9 in Ox), “cellular transport” (25 in Cx/8 in Ox), “protein synthesis” (18 in Cx/12 in Ox), “amino acid metabolism” (26 in Cx/2 in Ox), “secondary metabolism” (24 in Cx/2 in Ox), “cell rescue, defense and virulence” (11 in Cx/15 in Ox) “protein fate” (14 in Cx/11 in Ox), “biogenesis of cell wall components” (13 in Cx/8 in Ox), “energy” (17 in Cx/0 in Ox), “biogenesis of cytoskeleton” (10 in Cx/2 in Ox), and “cellular communication/signal transduction mechanism” (8 in Cx/3 in Ox).

The functional categories with larger numbers of up-regulated genes in Cx compared to Ox are consistent with structural and chemical modifications of the cell wall, and

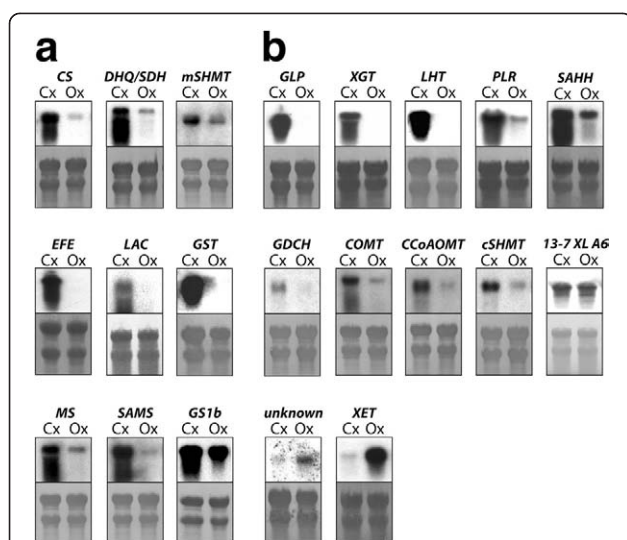


Figure 4 Validation of the expression profiles of selected genes in compression and opposite developing xylem. Equal amounts of total RNA isolated from compression (Cx) and opposite (Ox) differentiating xylem from adult maritime pine trees were size fractionated by electrophoresis and transferred to nylon membranes. PCR-amplified cDNA from 21 genes identified as differentially expressed in analyses of cDNA microarray 1 (a) or cDNA microarray 2 (b) were labeled with ^{32}P and used as probes. The panel 13-7 XL A6 shows the result of hybridization with cDNA from a gene classified as non-differentially expressed by the microarray analyses. Panels below every northern blot show methylene blue staining, as a loading control. Functional annotation and accession number of the corresponding ESTs: CS (chorismate synthase, BX250851), DHQ/SDH (3-dehydroquinate dehydratase/shikimate 5-dehydrogenase, BX251808), mSHMT (mitochondrial serine hydroxymethyltransferase, BX249820), EFE (ethylene-forming enzyme, BX251740), LAC (laccase, BX252150), GST (glutathione S-transferase, BX250376), MS (methionine synthase, BX255477), SAMS (S-adenosylmethionine synthetase, BX248790), GS1b (glutamine synthetase 1b, BX253698), GLP (germin-like protein, FN256510), XGT (xyloglucan galactosyltransferase, FN256497), LHT (lysine/histidine transporter, FN256535), PLR (pinorensinol-lariciresinol reductase, FN256693), GDCH (glycine decarboxylase complex H-protein, FN256878), COMT (caffeate O-methyltransferase, FN256925), CCoAOMT (caffeoyl-CoA-O-methyltransferase, FN256577), cSHMT (cytosolic serine hydroxymethyltransferase, FN256831), SAHH (S-adenosyl homocysteine hydrolase, FN256791), unknown (unknown-function gene, FN256975), XET (xyloglucan endotransglucosylase/hydrolase, FN256785), 13-7XL A6 (cDNA of a gene with the same level of expression in Cx and Ox according to microarray analyses).

as shown in Additional file 3 (“Up-regulated in Cx” tab) include genes encoding cellulose synthase subunits (contigs 19 and 7S, singletons BX249614 and FN256963), sucrose synthase (contigs 12 and 35S), cellulose synthase-like A (contig 19S), glycosyltransferases (contig 4P, singleton BX250177), xyloglucan galactosyltransferase (singleton FN256497), α - and β -tubulins (contigs 13, 7P, 31S and 51S, singletons BX255503, FN256632, FN256674, FN256784, BX255285, BX249611 and FN256695) and a putative microtubule-associated protein (singleton BX253390). In

particular, the gene encoding xyloglucan galactosyltransferase (XGT) showed a high level of expression in Cx, whereas transcripts were undetectable in Ox (Figure 4b, XGT).

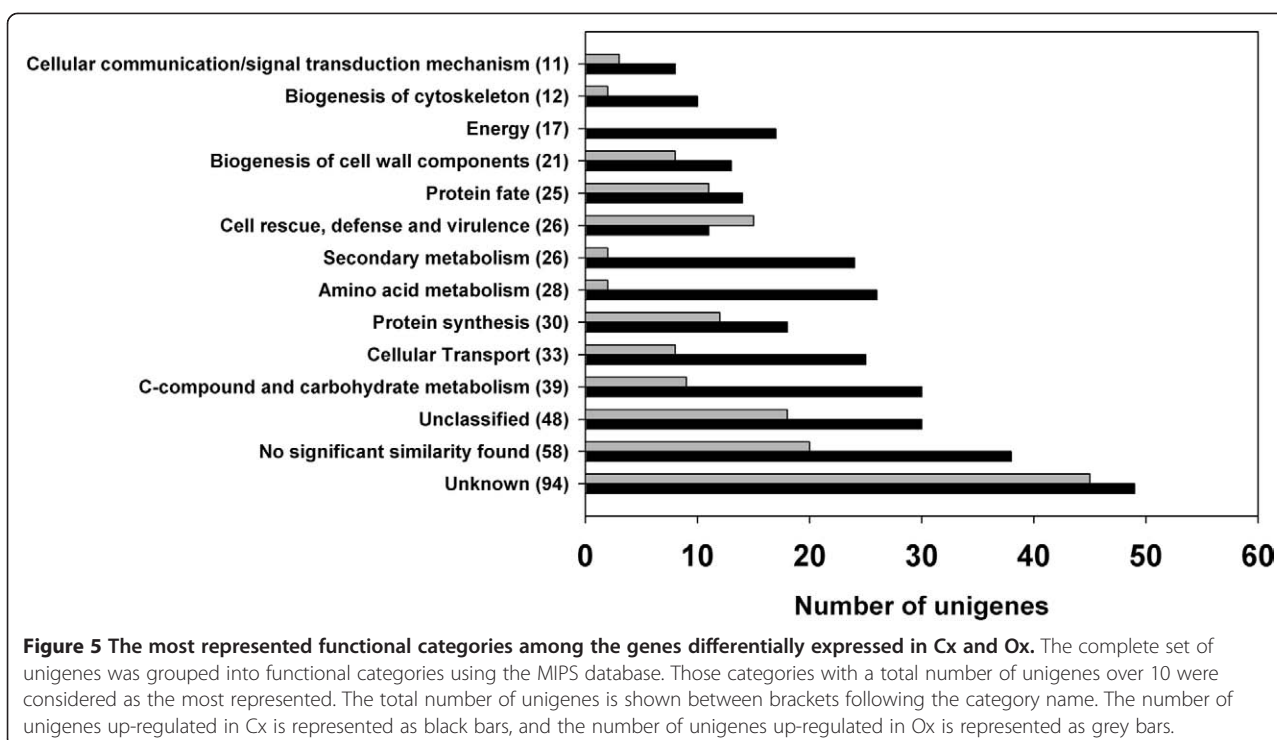
Transcripts for enzymes involved in lignin biosynthesis were also up-regulated in Cx (Additional file 3, “Up-regulated in Cx” tab), including those encoding enzymes in the monolignol biosynthetic pathway, such as phenylalanine ammonia lyase (PAL: contig 3S, singletons BX249569, BX253670 and FN257113), p-coumarate 3-hydroxylase (C3H: singleton BX248886), 4-coumarate:CoA ligase (4CL: singletons BX255483 and BX253310), cinnamyl alcohol dehydrogenase (CAD: contig 2P, singleton BX253572), hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase (HCT, singleton FN256636), caffeoyl-CoA-O-methyltransferase (CCoAOMT, singleton FN256577) and caffeate O-methyltransferase (COMT, contig 19P).

Several genes up-regulated in Cx and classified in the functional category “amino acid metabolism” encode enzymes involved in the synthesis of precursors for monolignol biosynthesis, such as the enzymes of the main trunk of the shikimate pathway chorismate synthase (contig 5, singleton FN256628), 3-dehydroquinate dehydratase/shikimate 5-dehydrogenase (contig 22), shikimate kinase (singleton BX251461) and 5-enolpyruvylshikimate 3-phosphate synthase (singleton BX252482). In particular, the expression pattern was confirmed by northern blot for the genes encoding chorismate synthase, and 3-dehydroquinate dehydratase/shikimate 5-dehydrogenase (Figure 4a, CS and DHQ/SDH, respectively).

Expression of genes encoding enzymes related to SAM biosynthesis in cells undergoing xylogenesis

Genes related to S-adenosylmethionine (SAM) metabolism were also widely represented among those up-regulated in Cx and classified in the functional category “amino acid metabolism” (Additional file 3, “Up-regulated in Cx” tab). SAM is a universal methyl donor for many different cellular metabolic reactions, including those catalyzed by CCoAOMT and COMT. Both enzymes are essential for the biosynthesis of coniferyl and sinapyl alcohols, precursors of lignin, lignans and other phenylpropanoids [34]. Therefore, the higher lignin content of compression wood may require a larger supply of SAM during development compared to opposite wood.

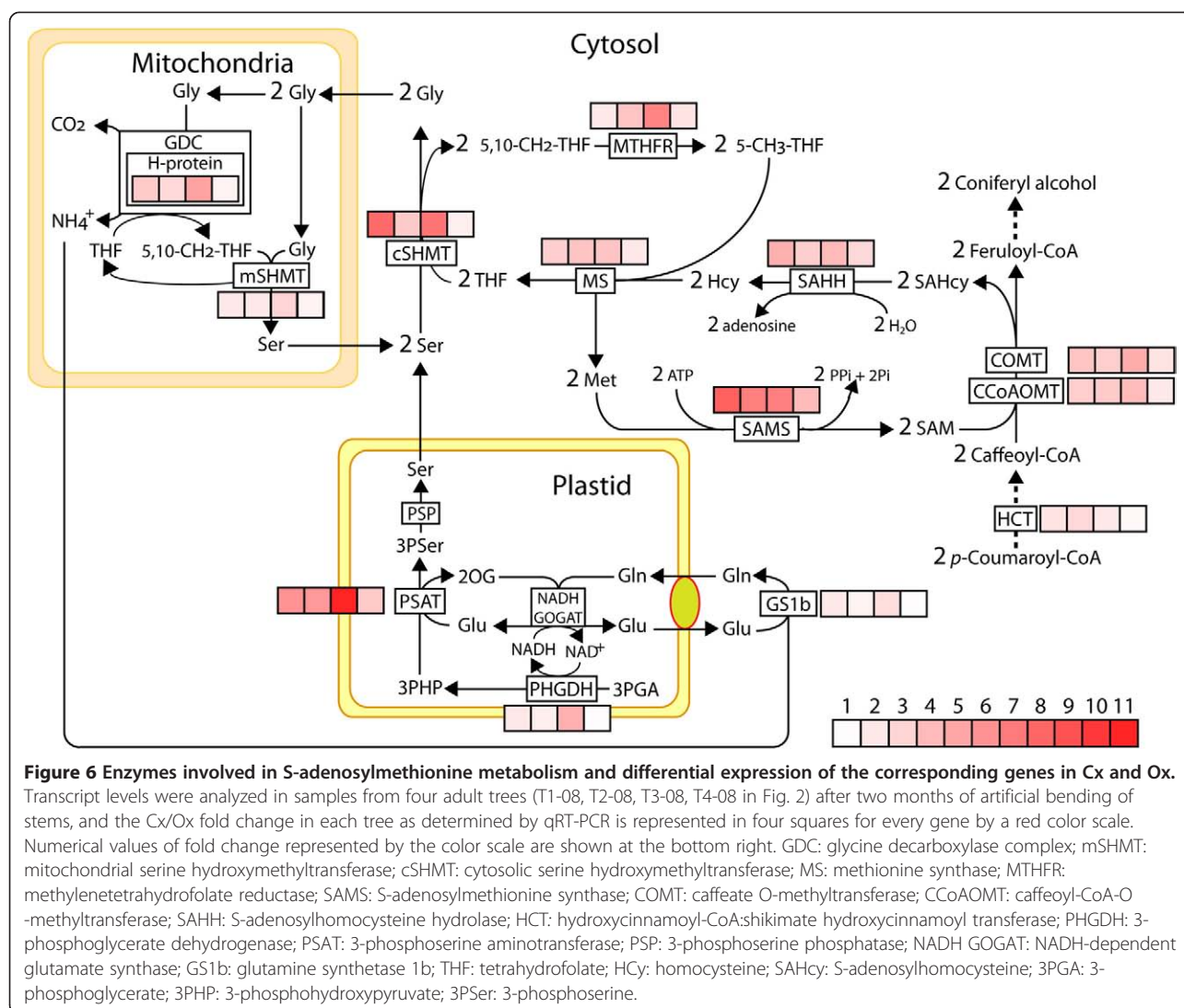
To unambiguously validate the differential expression patterns observed in the microarray analyses for genes related to SAM metabolism, relative transcript levels were analyzed by qRT-PCR in Cx and Ox of four independent trees from Sierra Bermeja that were sampled in 2008 (Figure 6 shows fold change values of relative transcript abundance and values for normalized relative abundance are provided in Additional file 4). These genes encode the enzymes of the activated methyl cycle and included methionine synthase (MS: contigs 27 and 21S, singletons BX255406, BX255477, BX251773, FN257093 and



FN256480), S-adenosylmethionine synthase (SAMS: singletons BX248790, FN257080, FN256463, FN256552, FN564393) and S-adenosylhomocysteine hydrolase (SAHH: contig 22S). The expression of genes encoding enzymes of the monolignol pathway branch that utilizes methyl groups provided by SAM to produce methoxylated monolignols were also analyzed, including hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase (HCT), Caffeoyl-CoA-O-methyltransferase (CCoAOMT) and Caffeate O-methyltransferase (COMT). Other genes included in the analysis encode enzymes involved in the supply of methyl units into the activated methyl cycle from the amino acid serine, such as cytosolic methylenetetrahydrofolate reductase (MTHFR: singletons BX254818 and FN564392) and serine hydroxymethyltransferase (cSHMT: contig 11 and singleton FN256755), mitochondrial glycine decarboxylase complex H-protein (GDCH: singleton FN256878) and serine hydroxymethyltransferase (mSHMT: singleton BX249820), and the plastid enzymes 3-phosphoglycerate dehydrogenase (PHGDH: singletons BX251805, FN256792) and phosphoserine aminotransferase (PSAT: singleton FN257119). Finally, we also considered a gene encoding glutamine synthetase (GS1b: singleton BX253698) up-regulated in Cx according to microarray analysis (Additional file 3), which has been proposed to play a key role in re-assimilation of ammonium released from different process in developing xylem [19,35]. To unambiguously verify the identity of the encoded proteins, we cloned by RLM-RACE the full-length cDNAs for

those genes that were not available in databases (see Additional file 5 for accession numbers). All these genes were indeed up-regulated in Cx compared to Ox in the four trees according to fold change values of relative transcript abundance (Figure 6 and Additional file 4). However, it is interesting to note that smaller differences in transcript levels between the two types of xylem were observed in tree T4-08. As a control, the relative transcript levels for two reference genes used in the normalization of qRT-PCR data are shown in Additional file 6.

To further demonstrate the relationship between lignin accumulation and enhanced expression of genes related to SAM metabolism, we quantified by qRT-PCR the relative levels of transcripts in cells from *Pinus pinaster* calli that were stimulated to differentiate tracheids *in vitro*. We tested the occurrence of tracheids after 0, 7 and 15 days of transfer to the induction medium (EDM without hormones) by phloroglucinol staining and microscopy inspection (Figure 7a). Differentiated tracheids were present even in the control medium (EDM with hormones), but the number of differentiated tracheids was higher in induction medium (data not shown). Consequently lignin levels were higher in samples from induction medium (Figure 7b). Transcript levels of genes encoding enzymes involved in SAM metabolism were higher in cells transferred to induction medium compared to those transferred to control medium, in parallel with the observed increase in lignin (Figure 7c). In most



of the genes the level of transcripts in cells after 15 days in induction medium was at least twice the level in cells that were cultured in control medium.

Expression of genes encoding enzymes related to lignan-biosynthesis in cells undergoing xylogenesis

In addition to being a lignin precursor, coniferyl alcohol is also a precursor of different lignans [36]. Three genes encoding enzymes involved in biosynthesis of coniferyl alcohol-derived lignans were also found among the up-regulated genes in Cx, according to microarray analyses (Additional file 3). Pinoresinol-lariciresinol reductase (PLR: contig 44S, singletons BX251658 and BX252917) is an enzyme involved in the 8-8' linked lignans pathway, and phenylcoumaran benzylic ether reductase (PCBER: contigs 17 and 17S, singleton FN256729) and phenylpropanal double-bond reductase (PPDBR: contig 36S,

singleton BX255256) are involved in the 8-5' linked lignans pathway [37]. The differential expression of genes encoding these enzymes was also validated in four trees by qRT-PCR, with higher transcript levels in Cx compared to Ox (Figure 8a). Notably, this difference was also lower in tree T4-08. Transcript levels were also increased in cells of *Pinus pinaster* calli that were induced to differentiate into tracheids compared to non-induced cells (Figure 8b).

Discussion

Comparative transcriptome analysis shows reprogramming of cell wall related genes during compression wood formation in pine

In this work, we carried out a comparative transcriptome analysis between developing compression and opposite wood using a combination of differently constructed cDNA

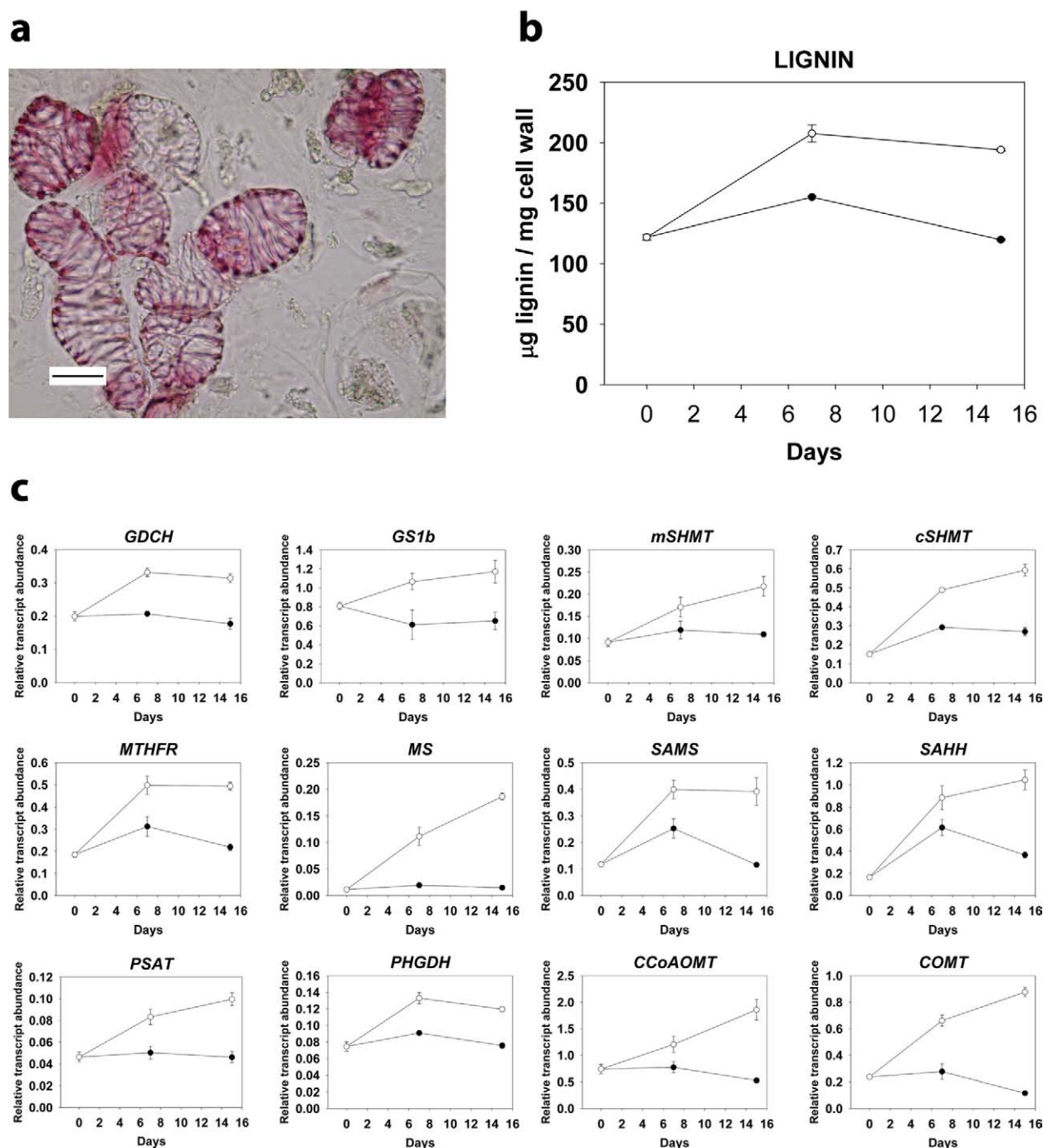


Figure 7 Lignin accumulation and expression of genes involved in S-adenosylmethionine metabolism during *in vitro* tracheid differentiation from *Pinus pinaster* callus cells. (a) Microscopic image of *Pinus pinaster* tracheids developed *in vitro*. Lignin deposited in a typical band pattern was stained with phloroglucinol. Scale bar on the bottom left represents 10 μ m. (b) Lignin contents in culture of *Pinus pinaster* cells after 0, 7 and 15 days of transfer to induction medium (without hormones, open circles) or control medium (with hormones, closed circles). Values are the means of three independent quantifications \pm SE. (c) Relative transcript abundance of genes involved in S-adenosylmethionine metabolism after 0, 7 and 15 days of transferring *Pinus pinaster* callus cells to induction medium (without hormones, open circles) or control medium (with hormones, closed circles). Values are the means of three independent replicates \pm SE.

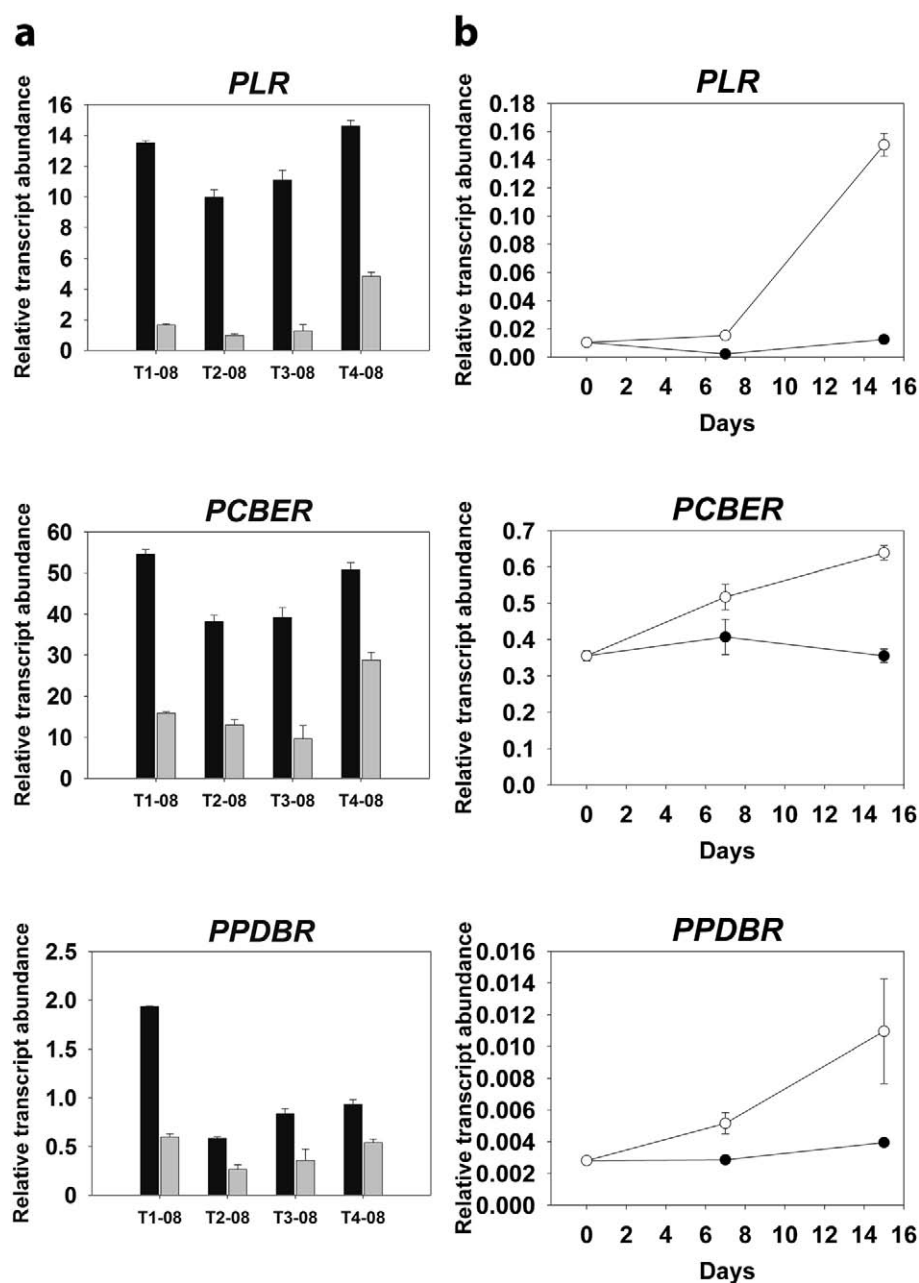


Figure 8 Relative transcript abundance of three genes involved in lignan biosynthesis in cells undergoing xylogenesis. Relative transcript abundance for each gene was determined by qRT-PCR. (a) Relative transcript levels in compression (black bars) and opposite (grey bars) developing xylem from four adult trees (T1-08, T2-08, T3-08, T4-08) after two months of artificial bending of the stems. (b) Relative transcript levels after 0, 7 and 15 days of transferring *Pinus pinaster* callus cells to induction medium (without hormones, open circles) or control medium (with hormones, closed circles). Values are means \pm SE of three independent replicates. *PLR*: Pinoresinol-lariciresinol reductase; *PCBER*: phenylcoumaran benzylic ether reductase; *PPDBR*: phenylpropenal double-bond reductase.

libraries and microarray analyses (Figure 1). Samples from three unrelated provenances, different geographic location or collected in different years (Figure 2) were integrated in the analysis to mitigate the potential effects of variability triggered by genotypic or environmental differences in order to identify a consistent set of candidate genes for wood properties.

The external stimulus that leads to the formation of compression wood caused significant changes of gene expression, as revealed by the large number of genes that were up-regulated or down-regulated in Cx compared to Ox (Figure 3) and the specific expression of some genes in Cx (Figure 4). An important component of this gene re-programming is clearly related to the chemical and

structural modifications of the thick secondary cell wall characteristic of conifer reaction wood, as suggested by the over-represented functional categories among the up-regulated genes in Cx, such as biogenesis of cytoskeleton, biogenesis of cell wall components, secondary metabolism, cellular transport and C-compound and carbohydrate metabolism (Figure 5 and Additional file 3).

A parallel up-regulation of genes encoding cellulose synthase subunits and sucrose synthase in Cx compared to Ox was observed, suggesting a higher rate of cellulose biosynthesis during compression wood formation. It has been well documented that the relative content of cellulose in compression wood is lower than in normal wood and opposite wood [21,38]. However, the S2 layer of secondary cell walls in compression wood is thicker than in normal and opposite wood, which contributes to the final thickness of the tracheid cell wall [21]. Therefore, in terms of absolute quantity, the development of compression tracheids may require higher rates of cellulose synthesis.

Different genes encoding α - and β -tubulins were also up-regulated during compression wood formation. Up-regulation of α -tubulin during compression wood formation in *Pinus taeda* has been reported previously [16], which supports our findings. Different lines of experimental evidence suggest that cortical microtubules of the cytoskeleton are involved in secondary cell wall pattern deposition and cellulose microfibril orientation [39,40]. We also identified a gene up-regulated in Cx that encodes a putative microtubule-associated protein (MAP). Several MAPs have been investigated [41] and their involvement in regulating cellulose microfibril angle (MFA) has been revealed in *Arabidopsis* MAP mutants, which show altered MFA in fiber walls [42,43]. More recently, it has been demonstrated that microtubule-associated protein AtMAP70-5 regulates secondary cell wall patterning in *Arabidopsis* [44]. Differential expression of these genes may contribute to the remarkable variation of MFA in the S2 cell wall layer between compression and opposite or normal wood, which plays a key role in the different mechanical properties of both types of wood [20,21,29,45].

Another group of genes that are up-regulated during compression wood differentiation encode enzymes that could be involved in the biosynthesis of hemicellulose, including a cellulose synthase-like, two glycosyltransferases and a xyloglucan galactosyltransferase. Although xyloglucan is a hemicellulose characteristic of primary cell walls, it has been suggested that modifications in xyloglucan may be related to the reinforcement of connections between primary and secondary wall layers [46]. In particular, the gene encoding a putative xyloglucan galactosyltransferase seems to be specifically induced during Cx formation (Figure 4b, *XGT* and data not shown). In contrast, two genes encoding a putative

xyloglucan endotransglucosylase/hydrolase are down-regulated in Cx compared to Ox (Additional file 3, "Up-regulated in Ox" tab and Figure 4b, *XET*). Therefore, our comparative transcriptome analysis suggests specific modifications of hemicellulose biosynthesis during compression wood development.

Consistent with the enrichment of lignin in compression wood [21], we found that genes encoding enzymes of the monolignol biosynthetic pathway were up-regulated in Cx (Additional file 3, "Up-regulated in Cx" tab), such as phenylalanine ammonia-lyase (PAL), 4-coumarate:CoA ligase (4CL), p-coumarate-3-hydroxylase (C3H), cinnamyl-alcohol dehydrogenase (CAD) hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase (HCT), caffeoyl-CoA-O-methyltransferase (CCoAOMT) and caffeate O-methyltransferase (COMT). In addition, the biosynthesis of monolignols demands substrates such as phenylalanine, shikimate and S-adenosylmethionine (SAM), which are provided by different metabolic reactions [12]. The quantitative importance of the shikimate and SAM metabolic pathways in the context of a highly lignified tissue, such as compression xylem, is supported by the fact that 22 of the 26 genes up-regulated in Cx that were classified in the "amino acid metabolism" functional category encode enzymes in these pathways (Additional file 3, "Up-regulated in Cx" tab).

In agreement with the elevated levels of shikimate in compression wood compared to normal wood that has been reported [30] and the higher demand of this substrate for monolignol biosynthesis, our microarray analyses showed that four genes encoding enzymes in the main trunk of the shikimate pathway were up-regulated in Cx (chorismate synthase, 3-dehydroquinate dehydratase/shikimate 5-dehydrogenase, shikimate kinase and 5-enolpyruvylshikimate 3-phosphate synthase).

Transcriptional adaptation of SAM biosynthetic pathway to the metabolic demand of coniferyl alcohol during xylogenesis

SAM is a methyl-group donor for coniferyl alcohol biosynthesis in conifers, and for both coniferyl and sinapyl alcohol in angiosperms [34]. Genes involved in lignin biosynthesis and those encoding enzymes of the activated methyl cycle were shown to be more highly expressed in the woody core tissue of hemp [47]. According to this functional relationship, genes encoding the enzymes of the activated methyl cycle (MS, SAMS and SAHH) were also up-regulated in Cx (Additional file 3 and Figure 6). Moreover, six additional genes were also up-regulated in Cx that encode enzymes located in three different cellular compartments, which are involved in providing a continuous supply of methyl groups from serine into the cycle (Figure 6).

Two of these genes encode cytosolic serine hydroxymethyltransferase (cSHMT) and methylenetetrahydrofolate reductase (MTHFR), which catalyse the conversion of serine to glycine with the concomitant transfer of a methyl group to the tetrahydrofolate to produce 5,10-methylenetetrahydrofolate and the reduction of 5,10-methylenetetrahydrofolate (CH₂-THF) to 5-methyltetrahydrofolate (CH₃-THF) respectively (Figure 6). Genes encoding H-protein subunit of the mitochondrial glycine decarboxylase complex (GDCH) and the mitochondrial serine hydroxymethyltransferase (mSHMT) are also included in this functional group. The essential function of the mitochondrial glycine decarboxylase/serine hydroxyl methyltransferase complex in C1 metabolism has been previously described [48]. Glycine catabolism and recycling to serine by GDC/SHMT complex prevents the accumulation of glycine under high rates of SAM synthesis. The accumulation of glycine would push the cSHMT reaction toward the formation of serine and inhibit SAM synthesis [48]. H-protein is a critical element in the sequence of reactions in the GDC complex, acting as a mobile substrate for the other enzymatic components, and changes in the levels of H-protein may modulate the activity of the complex [49]. Different genes encoding H-protein have been found in *Populus*, and transcripts encoding individual isoforms were specifically abundant in developing xylem, suggesting the association of distinct isoforms with photorespiration and C1 metabolism [50]. Finally, genes encoding plastid enzymes 3-phosphoglycerate dehydrogenase (PGDH) and phosphoserine aminotransferase (PSAT) are involved in serine biosynthesis from 3-phosphoglycerate [51], which would supply the high demand of this amino acid that is required for SAM synthesis during lignification.

The coordinated modulation of SAM and monolignol biosynthetic genes with lignin accumulation was also supported by two additional observations. First, the smallest differences in transcript abundances between Cx and Ox were observed for tree T4-08 (Figure 6), which also showed the smallest differences in lignin content (Figure 2b). Second, the relative transcript abundance of these genes during *in vitro* tracheid differentiation was parallel to lignin content (Figure 7). The increase in the levels of all enzymes involved in SAM synthesis may allow an increased flux through the pathway to supply the massive demand of methyl groups during lignification, with the minimal impact on metabolite levels [52]. In conclusion, these data strongly suggest that up-regulation of SAM metabolism genes during compression wood formation is due to an increased demand for methyl groups in coniferyl alcohol biosynthesis. However, SAM also acts as the precursor in the biosynthesis of the polyamines spermidine and spermine, the metal ion chelating compounds nicotianamine and the plant hormone ethylene. When acting as a precursor of these molecules, SAM is recycled from released 5'-

methylthioadenosine through the Yang cycle [53], but genes encoding enzymes of this pathway were not found among those up-regulated in Cx.

As a result of the activity of mitochondrial GDC during lignification the release of significant amounts of ammonium should be expected (Figure 6), in addition to the ammonium released in the reaction catalysed by PAL. The massive ammonium release could compromise nitrogen economy in a woody perennial during secondary growth if it is not efficiently recycled [19]. It appears that in order to cope with the high amounts of ammonia released by PAL and GDC activities during lignification a glutamine synthetase gene (*GS1b*) was up-regulated in Cx and during *in vitro* tracheid differentiation (Figure 6 and Figure 7). This gene has been functionally associated with ammonia assimilation in vascular tissues of seedlings [35,54].

In conifers, lignin lacks sinapyl alcohol, and coniferyl alcohol is the most abundant monolignol [12] in the polymer. Compression wood is relatively enriched in H subunits, and an increase in p-glucocoumaryl alcohol is associated with development of this reaction wood [30]. Nevertheless, the increased level of transcripts for enzymes involved in SAM synthesis in Cx indicates a higher demand for coniferyl alcohol to respond to an overall increase in lignin production, which is consistent with the reported higher abundance of coniferin residues in compression wood relative to normal wood in *Pinus taeda* [30]. However, an additional fate for coniferyl alcohol is the synthesis of lignans [37]. Three genes encoding enzymes involved in coniferyl alcohol-derived lignan metabolism were also up-regulated in Cx compared with Ox and in cells that were induced to differentiate into tracheids *in vitro* compared to non-induced cells (Additional file 3, Figure 8): pinoresinol-lariciresinol reductase (PLR), phenylcoumaran benzylic ether reductase (PCBER) and phenylpropanal double-bond reductase (PPDBR). These results suggest that during compression wood formation an increased demand for coniferyl alcohol residues may also be related to an increase in the biosynthesis of lignans. Lignans have been related to defense and heartwood formation [36]. Moreover, PCBER is a prominent poplar xylem protein that is strongly associated with lignifying cells [55], and a general function for PCBER has been proposed in lignifying tissues and wood development. Lignans may be infused in the secondary cell wall during lignification to cope with the oxidative stress accompanying lignin polymerization [55], which may explain the enhanced expression of lignan biosynthesis genes during compression wood formation and *in vitro* tracheid differentiation.

Conclusions

Our strategy for comparative transcriptomic analysis in pine Cx and Ox has revealed a broad spectrum of changes

in the transcriptome of differentiating xylem cells, consistent with the formation of wood with different structures and compositions. This work provides a resource for further characterization of molecular components that determine variation in softwood properties. In particular, the coordinated modulation of expression that was observed for genes involved in serine, SAM, monolignol and lignan biosynthesis, and the central role of glutamine synthetase to avoid N deficiency are important for understanding plant metabolic requirements and regulation during intensive lignification, and should thus be considered in strategies for lignin bioengineering.

Methods

Plant material

To construct the different cDNA libraries for microarray manufacturing, *P. pinaster* samples of compression (Cx), opposite (Ox), early (Ex) and late (Lx) developing xylem were collected from different genotypes of a Corsican clonal population planted in 1986 in the forestry station of INRA-Pierroton (Aquitaine, France). Compression wood was induced by bending the stem at an angle of 15° away from the vertical, and 12 samples were collected at the beginning (early wood) and at the end (late wood) of the growing season in 1998 and 1999 (after 8 days, 40 days, 120 days and 1.5 year of starting the stimulus) and 2000 (after 6 hours and 1 days after starting the stimulus). Cx was sampled from the underside of the bent trunk and Ox from the upper side. Ex and Lx samples were collected from a 14-year-old straight tree in April and August 1999, respectively. Juvenile (Jx) and Mature (Mx) developing xylem were sampled in 2001 from the crown and the base (breast height) of the stem in a 30-year-old maritime pine tree from the Aquitaine provenance.

For microarray and northern hybridizations, samples of Cx and Ox were collected from four maritime pine trees between 25 and 35 years old (T1-05, T2-05, T3-05 and T4-05) in May 2005 in Sierra Bermeja (Estepona, Spain). For real-time quantitative PCR analyses (qRT-PCR), samples of Cx and Ox were collected in May 2008 in Sierra Bermeja from four different trees (T1-08, T2-08, T3-08 and T4-08). Cx was induced by bending the tree stem at an angle of 45° for 60 days before sampling. Developing xylem was scraped with a scalpel after removing bark and phloem. Samples were immediately frozen in liquid nitrogen after harvesting and stored at -80 °C until use.

Development of *Pinus pinaster* calli and *in vitro* differentiation of tracheids

Calli were developed from *Pinus pinaster* hypocotyl explants using the procedures and culture media described by Möller *et al.* [56]. To induce tracheids, callus cells were suspended in EDM liquid medium and transferred to

induction media following the same procedure as described by Möller *et al.* [56], except that the induction medium was EDM medium without hormones instead of EDM medium supplemented with activated charcoal. EDM medium with hormones was inoculated in parallel as a control medium.

Tissue preparation and phloroglucinol staining for light microscopy

Xylem scrapings were fixed by the freeze substitution method [57]. Tissue pieces were embedded in Paraplast (Leica Microsystems) at 42 °C in a progression series with Histo-Clear and then incubated six times in Paraplast at 62 °C for 8 h. Sections (10 µm) were then obtained from embedded tissues with a Leitz microtome (Ernst Leitz, Midland, Ontario, Canada) and directly mounted onto poly-L-Lys-coated glass slides. Sections were stained with phloroglucinol (Sigma-Aldrich) following the procedure described elsewhere [58] and visualized under light microscopy, using a Nikon Eclipse E 800 microscope.

Lignin quantification

Cell wall preparation and lignin quantification was performed following the method described by Lange *et al.* [59]. For cell wall preparation 100 mg of tissue was used.

Isolation of RNA

Total RNA was isolated following the method of Chang *et al.* [60]. RNA concentration and purity was determined by spectrophotometry and integrity was confirmed by electrophoresis on denaturing agarose/formaldehyde gels.

Construction of cDNA libraries

The cDNA probes for microarray construction were obtained from two different sources. A composite cDNA library was constructed from a mixed pool of equal amounts of total RNA extracted from samples of the Corsican provenance described above, including Ex, Lx and the Cx and Ox samples from artificially bent trees for different periods of time (from 6 hours to 1.5 years). PolyA(+) RNA was isolated from the mixed pool of total RNA, and the cDNA library was constructed using the λ-ZAP-cDNA synthesis kit (Stratagene, La Jolla, CA, USA).

The second source of cDNA probes was four subtractive cDNA libraries constructed from samples of the Corsican provenance by the suppression subtractive hybridization method (SSH), using the PCR-Select cDNA Subtraction Kit (Clontech, Palo Alto, CA, USA). Two SSH libraries were produced using cDNAs from Cx and Ox by performing the subtraction procedure in both directions. The same procedure was followed to obtain two SSH libraries with samples of Jx and Mx. Subtracted cDNAs were cloned in pGEM®-T Easy plasmid (Promega, Madison, WI, USA), transformed into *Escherichia coli* JM109 (Promega) or XL1Blue (Stratagene, La Jolla, CA, USA) and

plated on LB plates containing 100 µg/mL ampicillin, 1 mM IPTG and 80 µg/mL X-gal. White colonies were selected from each library and separately cultured in 96-well plates containing LB with ampicillin.

Construction of cDNA microarrays

To construct cDNA microarray 1, 2800 inserts from the composite λ-ZAP cDNA library were PCR-amplified using T3 and T7 primers or M13 Forward and M13 Reverse primers. PCR products were purified with 96-well multi-screen filter plates (Millipore Corp. Bedford, MA, USA). Amplified cDNAs were checked on agarose gel electrophoresis after purification. Solutions of purified-cDNA probes in 50% DMSO were prepared to a concentration between 100 and 200 ng/µl, and stored at -20 °C until use. Each cDNA probe was printed in duplicate onto ULTRA gaps II coated slides (Corning Inc., NY, USA) using a Qarray2 (Genetix Ltd, Queensway, UK) with a telechem printing head and 16 split pins (Biorobotics, Cambridge, UK), and with a 4 x 4 configuration. As a control, ArrayControl Sense Oligo Spots (spikes) (Ambion Inc., Austin, TX, USA) were included. After printing, the slides were dried at room temperature and the spotted cDNA were cross-linked to the slide surface by UV irradiation at 300 mJ/cm².

For cDNA microarray 2, 4041 cDNA clones from the four SSH libraries described above were randomly selected and cDNAs inserts were amplified by PCR using Nested PCR primers 1 and 2R included in the PCR-select cDNA subtraction kit (Clontech, Palo Alto, CA, USA). PCR-amplified cDNA were purified and spotted in duplicate as described above.

Microarray hybridization, scanning and data acquisition

Dye labeled aRNA was synthesized with the Amino allyl MessageAmp II aRNA amplification kit (Ambion Inc., Austin, TX, USA). Prehybridization, hybridization, and posthybridization washes were carried out using Pronto™ Universal Hybridization Kit (Corning Inc.) in a HS 400 Pro hybridization station (Tecan Trading AG, Switzerland). Five micrograms of each of the labeled aRNA targets was added to the hybridization solution and denatured at 95 °C for 5 min. The slides were hybridized for 16 h at 42 °C. Hybridized slides were scanned with 5 µm resolution, and signal intensities were detected with a Q-Scan scanner (Genetix).

Microarray data analysis

Four dual-target hybridizations were performed using labeled aRNA targets from Cx and Ox samples from different individual pine trees (T1-05, T2-05, T3-05 and T4-05), including two dye-swap experiments. Since every slide was printed with two full replicates of the microarray, each microarray data set consisted of four dye-balanced hybridizations for

each type of xylem in duplicate. Spots flagged below 0 using GenePix v6.0 software (Axon Instruments) and those with signal intensities that did not surpass 2X the background signal in both channels were discarded. Background correction was performed with the “normexp” method of the limma library. The M value was defined as the base two logarithm of every expression ratio, computed as the ratio between the background-corrected foreground intensities of the Cy3 and Cy5 channels. Raw expression data were normalized for all sources of systematic variation with the print-tip loess [61], using the common assumption of considering the whole microarray expression as invariant. Scaling between arrays was not needed. Gene significance was then estimated using a robust linear model corrected by a moderated *t*-test (empirical Bayes) [62]. The multi-testing effect was corrected adjusting *p*-values by the Benjamini and Hochberg method [63]. A gene was considered significantly up- or down-regulated if it met these two criteria: (1) adjusted $p \leq 0.001$; and (2) fold change ≥ 1.5 . The Biobase v 2.0.1 of the Bioconductor package [64] was installed under R version 2.7.1 for all statistical analysis, mainly the limma v 2.14.5 [65] and marray v 1.18.0 libraries (<http://www.bioconductor.org/packages/release/bioc/html/marray.html>). The outputs of microarray analyses carried out in this study are available at Gene Expression Omnibus data repository (GEO accession numbers GSE37678 and GSE37736). All data are MIAME compliant.

Sequencing and sequence analysis

A total of 12,134 plasmid clones from the composite cDNA library were sequenced by single pass from the 5'-end with the T3 primer, which rendered 8429 ESTs. These sequences were pre-processed and assembled into unigenes before proceeding to functional annotation as described elsewhere [32]. Sequencing of cDNA clones from the SSH libraries from genes that were differentially expressed in Cx and Ox according to microarray 2 data were carried out using M13 forward, SP6 or T7 primers. Vector and adapter sequences were deleted using SeqTrim [66]. Sequences with an internal *RsaI* site were discarded and not considered for contig assembly, as they were supposed to contain fusions of two cDNA fragments. All the sequences were deposited in the EMBL Nucleotide Sequence Database (accession numbers are listed in Additional file 3).

Contigs were established using CAP3 [67]. Non-redundant cDNA sequences after EST assembly were annotated by BLASTX at GenBank and classified into functional groups using data reported in the literature and MIPS functional catalogue [68]. Only BLASTX hits with $E\text{-values} \leq 1 \times 10^{-10}$ and a minimum of 50% similarity were selected for annotation. Sequences showing weak similarity ($E\text{-value} > 1 \times 10^{-10}$) and those without significant similarity were compared by BLASTN algorithm with the Pine Gene

Index database (<http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=pine>). The resulting pine tentative contigs (TCs) or pine ESTs, which shared more than 90% of identity with each query sequence, were used as new queries for BLASTX at GenBank.

Northern blot analysis

Seven micrograms of total RNA from Cx or Ox were analyzed by Northern blot as previously described [69]. ³²P labeled probes of selected genes were produced with the High Prime System (Roche Diagnostics). Prehybridizations and hybridizations were performed at 65 °C. Hybridized membranes were washed with 1X SSC, 0.1% SDS and 0.1X SSC, 0.1% SDS at 65 °C.

Real-time quantitative PCR analysis (qRT-PCR)

RNA samples were treated with RNase-Free rDNase to eliminate genomic DNA contamination and then purified using NucleoSpin RNA clean-up (Macherey-Nagel, Düren, Germany). The quality of the treated RNA was checked by both gel electrophoresis and spectrophotometry. Reverse transcription using 1 µg of total RNA was carried out with iScript cDNA Synthesis Kit (Bio-Rad Laboratories, CA, USA).

Sense and antisense primers were designed for the specific amplification of selected genes. The sequences of the primers are shown in Additional file 5. PCR reactions were performed in a Stratagene MxPro 3000P Real-Time PCR System (Stratagene). Reactions were performed in 25 µL containing Quantimix easy master mix (Quantimix Easy SYG kit, BioTools B&M Labs, S.A., Madrid, Spain), 0.4 µM of each primer and 10 ng of cDNA (RNA equivalent). PCRs were performed by incubation 2 min at 95 °C followed by 40 cycles of 30 s at 95 °C, 30 s at specific annealing temperature and 15 s at 72 °C. Fluorescence was measured at the end of each extension step. Three technical replicates were analyzed for each sample. The specificity of each amplification reaction was verified by melting point analysis at the end of each experiment, and during protocol development by gel electrophoresis. In every assay negative controls were included with RNA not reverse transcribed from the different analyzed samples or without template to rule out contaminations due to remaining genomic DNA or sample manipulations.

The values of amplification efficiency, Ct and initial fluorescence (R_0) for every reaction were calculated with the Miner algorithm [70]. Relative expression levels were obtained from the ratio between R_0 of the target gene and a normalization factor. To determine the normalization factor, the relative transcript abundances of five genes with fold change value of 1 in the microarray analyses were also determined by qRT-PCR (see Additional file 5 for primer sequences) and analyzed with the geNorm algorithm [71] to select the combination of more stable genes in every

experiment. For Cx versus Ox comparison the geometric mean of R_0 obtained for ribosomal protein L34 and actin was used. For the analysis of relative transcript abundance during *in vitro* development of tracheids the geometric mean of R_0 values for the genes encoding 40 S ribosomal protein S27, Ubi-like protein and elongation factor 1 α was used.

Additional file

Additional file 1: Relative expression values obtained with the analysis of microarray 1 and statistical significance. Spots flagged as bad, not found or absent and those with a signal intensity that did not surpass 2X their background signal were discarded. The relative expression level is expressed as the binary logarithm of the Ox/Cx expression ratio (log2 Fold Change). For each probe the values for the moderated t-statistic (t), associated p-value (P-value) and the associated p-value after adjustment for multiple testing (Adjusted P-value) are shown.

Additional file 2: Relative expression values obtained with the analysis of microarray 2 and statistical significance.

Additional file 3: Sequences of annotated EST and contigs of genes differentially expressed during compression and opposite wood formation. The sequence with the higher match score in the GeneBank (GeneBank match) together with the BlastX outcome values (eValue, Score, Positives) or the Pine Gene Index database (PGI tentative contig) are included. Values of relative expression (log2 Fold Change (Ox/Cx)) according to microarray analysis and statistical significance are also shown.

Additional file 4: Relative transcript abundance of genes involved in S-adenosylmethionine metabolism in Cx and Ox. Values for normalized relative abundance \pm SE used to construct Figure 6.

Additional file 5: Sequences of primers used in qRT-PCR.

Additional file 6: Relative transcript abundance of the reference genes encoding ribosomal protein L34 and actin in the RNA samples from Cx (black bars) and Ox (grey bars). Values are means \pm SE of three independent replicates.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We are grateful to Remedios Crespillo and Dr. Mónica Pérez Alegre for their technical assistance. We also thank Noé Fernández Pozo for the help with the submission of the sequences to the EMBL database, and Dr. Kaisa Nieminen for proofreading the manuscript. This work was supported by grants from Ministerio de Educación y Ciencia (AGL2006-07360 and AGL2009-12139-C02-02) Spain, and research funds from Junta de Andalucía (BIO-114). DPV was supported by a predoctoral fellowship from Junta de Andalucía (Spain). SMD-M and E-SSS were supported by pre-doctoral fellowships from Ministerio de Educación y Ciencia (Spain).

Author details

¹Departamento de Biología Molecular y Bioquímica, Facultad de Ciencias, Universidad de Málaga, Campus Universitario de Teatinos, 29071, Málaga, Spain. ²Department of Plant Molecular Biology, University of Lausanne, CH-1015, Lausanne, Switzerland. ³Division of Glycoscience, School of Biotechnology, Royal Institute of Technology, AlbaNova University Centre, SE-10691, Stockholm, Sweden. ⁴Departamento de Fisiología Vegetal, Centro Hispano-Luso de Investigaciones Agrarias, Facultad de Biología, Universidad de Salamanca, C/Río Duero 12, 37185, Salamanca, Spain.

Authors' contributions

DPV, SMD-M, FMC and FRC designed and constructed the microarrays. DPV, SMD-M, E-SSS, RAC and SHEK performed the experiments. DO and FRC constructed the different cDNA libraries. RB helped in managing and arranging cDNA clones for EST generation. MGC performed the computational analyses of hybridized



microarrays. DPV and FRC analyzed the data and made biological and literature interpretations of the results. FMC and MGC discussed the data and revised the manuscript. FRC designed the experiments, wrote the grants that funded the project, directed the project and wrote the manuscript. All authors read and approved the final manuscript.

Received: 12 January 2012 Accepted: 29 June 2012

Published: 29 June 2012

References

- Turner S, Gallois P, Brown D: **Tracheary element differentiation.** *Annu Rev Plant Biol* 2007, **58**:407–433.
- Schrader J, Nilsson J, Mellerowicz E, Berglund A, Nilsson P, Hertzberg M, Sandberg G: **A high-resolution transcript profile across the wood-forming meristem of poplar identifies potential regulators of cambial stem cell identity.** *Plant Cell* 2004, **16**:2278–2292.
- Paiva JA, Garnier-Gere PH, Rodrigues JC, Alves A, Santos S, Graca J, Le Provost G, Chaumeil G, Da Silva-Perez D, Bosc A, Fevèreiro P, Plomion C: **Plasticity of maritime pine (*Pinus pinaster*) wood-forming tissues during a growing season.** *New Phytol* 2008, **179**:1080–1094.
- Plomion C, Leprovost G, Stokes A: **Wood formation in trees.** *Plant Physiol* 2001, **127**:1513–1523.
- Zhang Y, Sederoff RR, Allona I: **Differential expression of genes encoding cell wall proteins in vascular tissues from vertical and bent loblolly pine trees.** *Tree Physiol* 2000, **20**:457–466.
- Hertzberg M, Aspeborg H, Schrader J, Andersson A, Erlandsson R, Blomqvist K, Bhalarao R, Uhlén M, Teeri T, Lundberg J, Sundberg B, Nilsson P, Sandberg G: **A transcriptional roadmap to wood formation.** *Proc Natl Acad Sci USA* 2001, **98**:14732–14737.
- Andersson-Gunnera S, Mellerowicz E, Love J, Segerman B, Ohmiya Y, Coutinho P, Nilsson P, Henricsson B, Moritz T, Sundberg B: **Biosynthesis of cellulose-enriched tension wood in *Populus*: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis.** *Plant J* 2006, **45**:144–165.
- Cato S, McMillan L, Donaldson L, Richardson T, Echt C, Gardner R: **Wood formation from the base to the crown in *Pinus radiata*: gradients of tracheid wall thickness, wood density, radial growth rate and gene expression.** *Plant Mol Biol* 2006, **60**:565–581.
- Foucart C, Paux E, Ladouce N, San-Clemente H, Grima-Pettenati J, Sivadon P: **Transcript profiling of a xylem vs phloem cDNA subtractive library identifies new genes expressed during xylogenesis in *Eucalyptus*.** *New Phytol* 2006, **170**:739–752.
- Paiva JA, Garces M, Alves A, Garnier-Gere P, Rodrigues JC, Lalanne C, Porcon S, Le Provost G, Perez Dda S, Brach J, Frigerio JM, Claverol S, Barre A, Fevèreiro P, Plomion C: **Molecular and phenotypic profiling from the base to the crown in maritime pine wood-forming tissue.** *New Phytol* 2008, **178**:283–301.
- Pavy N, Boyle B, Nelson C, Paule C, Giguere I, Caron S, Parsons LS, Dallaire N, Bedon F, Berube H, Cooke J, Mackay J: **Identification of conserved core xylem gene sets: conifer cDNA microarray development, transcript profiling and computational analyses.** *New Phytol* 2008, **180**:766–786.
- Boerjan W, Ralph J, Baucher M: **Lignin biosynthesis.** *Annu Rev Plant Biol* 2003, **54**:519–546.
- Aspeborg H, Schrader J, Coutinho P, Stam M, Kallas A, Djerbi S, Nilsson P, Denman S, Amini B, Sterky F, Master E, Sandberg G, Mellerowicz E, Sundberg B, Henricsson B, Teeri T: **Carbohydrate-active enzymes involved in the secondary cell wall biogenesis in hybrid aspen.** *Plant Physiol* 2005, **137**:983–997.
- Vanholme R, Morreel K, Ralph J, Boerjan W: **Lignin engineering.** *Curr Opin Plant Biol* 2008, **11**:278–285.
- Vander Mijsbrugge K, Meyermans H, Van Montagu M, Bauw G, Boerjan W: **Wood formation in poplar: identification, characterization, and seasonal variation of xylem proteins.** *Planta* 2000, **210**:589–598.
- Whetten R, Sun Y-H, Zhang Y, Sederoff R: **Functional genomics and cell wall biosynthesis in loblolly pine.** *Plant Mol Biol* 2001, **47**:275–291.
- Shen B, Li C, Tarczynski M: **High free-methionine and decreased lignin content result from a mutation in the Arabidopsis S-adenosyl-L-methionine synthetase 3 gene.** *Plant J* 2002, **29**:371–380.
- Sánchez-Aguayo I, Rodríguez-Galán J, García R, Torreblanca J, Pardo J: **Salt stress enhances xylem development and expression of S-adenosyl-L-methionine synthase in lignifying tissues of tomato plants.** *Planta* 2004, **220**:278–285.
- Cantón F, Suárez M, Cánovas F: **Molecular aspects of nitrogen mobilization and recycling in trees.** *Photosynth Res* 2005, **83**:265–278.
- Du S, Yamamoto F: **An overview of the biology of reaction wood formation.** *J Integr Plant Biol* 2007, **49**:131–143.
- Timell T: *Compression wood in gymnosperms.* Heidelberg: Springer; 1986.
- Donaldson L, Grace J, Downes G: **Within-tree variation in anatomical properties of compression wood in radiata pine.** *IAWA Journal* 2004, **25**:253–271.
- Sedighi-Gilani M, Sunderland H, Navi P: **Microfibril angle non-uniformities within normal and compression wood tracheids.** *Wood Sci Technol* 2005, **39**:419–430.
- Önnerud H: **Lignin structures in normal and compression wood. Evaluation by thioacidolysis using ethanethiol and methanethiol.** *Holzforchung* 2003, **57**:377–384.
- Yeh T-F, Goldfarb B, Chang H-m, Peszlen I, Braun J, Kadla J: **Comparison of morphological and chemical properties between juvenile wood and compression wood of loblolly pine.** *Holzforchung* 2005, **59**:669–674.
- Allona I, Quinn M, Shoop E, Swope K, St Cyr S, Carlis J, Riedl J, Retzel E, Campbell MM, Sederoff R, Whetten RW: **Analysis of xylem formation in pine by cDNA sequencing.** *Proc Natl Acad Sci USA* 1998, **95**:9693–9698.
- Pavy N, Laroche J, Bousquet J, Mackay J: **Large-scale statistical analysis of secondary xylem ESTs in pine.** *Plant Mol Biol* 2005, **57**:203–224.
- Yamashita S, Yoshida M, Yamamoto H, Okuyama T: **Screening genes that change expression during compression wood formation in *Chamaecyparis obtusa*.** *Tree Physiol* 2008, **28**:1331–1340.
- Yeh T-F, Braun J, Goldfarb B, Chang H-m, Kadla J: **Morphological and chemical variations between juvenile wood, mature wood, and compression wood of loblolly pine (*Pinus taeda* L.).** *Holzforchung* 2006, **60**:1–8.
- Yeh TF, Morris CR, Goldfarb B, Chang HM, Kadla JF: **Utilization of polar metabolite profiling in the comparison of juvenile wood and compression wood in loblolly pine (*Pinus taeda*).** *Tree Physiol* 2006, **26**:1497–1503.
- Alía R, Martín S, Miguel JD, Galera R, Agúndez D, Gordo J, Salvador L, Catalán G, Gil L: *Las regiones de procedencia de *Pinus pinaster* Aiton.* Madrid: DGCONA; 1996.
- Cantón F, Provost GL, García V, Barré A, Frigério J, Paiva J, Fevèreiro P, Avila C, Mouret J, Daruvar AD, Cánovas F, Plomion C: **Transcriptome analysis of wood formation in maritime pine.** In *Sustainable forestry, wood products & biotechnology*. Edited by Espinel S, Barredo Y, Ritter E. Vitoria-Gasteiz, Spain: DFA-AFA Press; 2003:334–347.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**:403–410.
- Roje S: **S-Adenosyl-L-methionine: beyond the universal methyl group donor.** *Phytochemistry* 2006, **67**:1686–1698.
- Avila C, Suarez M, Gomez-Maldonado J, Canovas F: **Spatial and temporal expression of two cytosolic glutamine synthetase genes in Scots pine: functional implications on nitrogen metabolism during early stages of conifer development.** *Plant J* 2001, **25**:93–102.
- Davin L, Lewis N: **An historical perspective on lignan biosynthesis: monolignol, allylphenol and hydroxycinnamic acid coupling and downstream metabolism.** *Phytochem Rev* 2003, **2**:257–288.
- Kasahara H, Jiao Y, Bedgar D, Kim S-J, Patten A, Xia Z-Q, Davin L, Lewis N: ***Pinus taeda* phenylpropenal double-bond reductase: purification, cDNA cloning, heterologous expression in *Escherichia coli*, and subcellular localization in *P. taeda*.** *Phytochemistry* 2006, **67**:1765–1780.
- Lohrasebi H, Mabey W, Roy D: **Chemistry and pulping feasibility of compression wood in black spruce.** *J Wood Chem and Technol* 1999, **19**:13–25.
- Spokevicius AV, Southerton SG, MacMillan CP, Qiu D, Gan S, Tibbits JF, Moran GF, Bossinger G: **Beta-tubulin affects cellulose microfibril orientation in plant secondary fibre cell walls.** *Plant J* 2007, **51**:717–726.
- Wightman R, Turner SR: **The roles of the cytoskeleton during cellulose deposition at the secondary cell wall.** *Plant J* 2008, **54**:794–805.
- Hamada T: **Microtubule-associated proteins in higher plants.** *J Plant Res* 2007, **120**:79–98.
- Zhong R, Burk DH, Morrison WH 3rd, Ye ZH: **A kinesin-like protein is essential for oriented deposition of cellulose microfibrils and cell wall strength.** *Plant Cell* 2002, **14**:3101–3117.
- Burk DH, Ye ZH: **Alteration of oriented deposition of cellulose microfibrils by mutation of a katanin-like microtubule-severing protein.** *Plant Cell* 2002, **14**:2145–2160.
- Pesquet E, Korolev AV, Calder G, Lloyd CW: **The microtubule-associated protein AtMAP70-5 regulates secondary wall patterning in *Arabidopsis* wood cells.** *Curr Biol* 2010, **20**:744–749.

45. Treacy M, Evertsen J, Dhuhhain AN: *A comparison of mechanical and physical wood properties of a range of sitka spruce provenances*. Dublin: National Council for Forest Research and Development (COFORD); 2000.
46. Bourquin V, Nishikubo N, Abe H, Brumer H, Denman S, Eklund M, Christiennin M, Teeri T, Sundberg B, Mellerowicz E: **Xyloglucan endotransglycosylases have a function during the formation of secondary cell walls of vascular tissues**. *Plant Cell* 2002, **14**:3073–3088.
47. Broeck Hvd, Maliepaard C, Ebskamp M, Toonen MAJ, Koops AJ: **Differential expression of genes involved in C1 metabolism and lignin biosynthesis in wooden core and bast tissues of fibre hemp (*Cannabis sativa* L.)**. *Plant Sci* 2008, **174**:205–220.
48. Mouillon J-M, Aubert S, Bourguignon J, Gout E, Douce R, Rébeillé F: **Glycine and serine catabolism in non-photosynthetic higher plant cells: their role in C1 metabolism**. *Plant J* 1999, **20**:197–205.
49. Douce R, Bourguignon J, Neuburger M, Rébeillé F: **The glycine decarboxylase system: a fascinating complex**. *Trends Plant Sci* 2001, **6**:167–176.
50. Rajinikanth M, Harding S, Tsai C-J: **The glycine decarboxylase complex multienzyme family in Populus**. *J Exp Bot* 2007, **58**:1761–1770.
51. Ho C-L, Saito K: **Molecular biology of the plastidic phosphorylated serine biosynthetic pathway in *Arabidopsis thaliana***. *Amino Acids* 2001, **20**:243–259.
52. Fell D: **Increasing the flux in metabolic pathways: a metabolic control analysis perspective**. *Biotechnol Bioeng* 1998, **58**:121–124.
53. Miyazaki J, Yang S: **The methionine salvage pathway in relation to ethylene and polyamine biosynthesis**. *Physiol Plant* 1987, **69**:366–370.
54. Canovas FM, Avila C, Canton FR, Canas RA, de la Torre F: **Ammonium assimilation and amino acid metabolism in conifers**. *J Exp Bot* 2007, **58**:2307–2318.
55. Vander Mijnsbrugge K, Beeckman H, De Rycke R, Van Montagu M, Engler G, Boerjan W: **Phenylcoumaran benzylic ether reductase, a prominent poplar xylem protein, is strongly associated with phenylpropanoid biosynthesis in lignifying cells**. *Planta* 2000, **211**:502–509.
56. Möller R, McDonald A, Walter C, Harris P: **Cell differentiation, secondary cell-wall formation and transformation of callus tissue of *Pinus radiata* D. Don**. *Planta* 2003, **217**:736–747.
57. Regan S, Bourquin V, Tuominen H, Sundberg B: **Accurate and high resolution in situ hybridization analysis of gene expression in secondary stem tissues**. *Plant J* 1999, **19**:363–369.
58. Canas RA, de la Torre F, Canovas FM, Canton FR: **Coordination of PsAS1 and PsASPG expression controls timing of re-allocated N utilization in hypocotyls of pine seedlings**. *Planta* 2007, **225**:1205–1219.
59. Lange B, Lapiere C, Sandermann H: **Elicitor-induced spruce stress lignin**. *Plant Physiol* 1995, **108**:1277–1287.
60. Chang S, Puryear J, Cairney J: **A simple and efficient method for isolating RNA from pine tree**. *Plant Mol Biol Rep* 1993, **11**:113–116.
61. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, Speed TP: **Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation**. *Nucleic Acids Res* 2002, **30**:e15.
62. Smyth G: **Linear models and empirical bayes methods for assessing differential expression in microarray experiments**. *Stat Appl Genet Mol Biol* 2004, **3**:1–26.
63. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful approach to multiple testing**. *J R Statist Soc B* 1995, **57**:289–300.
64. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J: **Bioconductor: open software development for computational biology and bioinformatics**. *Genome Biol* 2004, **5**:R80.
65. Smyth G: **Limma: linear models for microarray data**. In *Bioinformatics and computational biology solutions using R and Bioconductor*. Edited by Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W. New York: Springer; 2005:397–420.
66. Falgueras J, Lara AJ, Fernandez-Pozo N, Canton FR, Perez-Trabado G, Claros MG: **SeqTrim: a high-throughput pipeline for pre-processing any type of sequence read**. *BMC Bioinformatics* 2010, **11**:38.
67. Huang X, Madan A: **CAP3: a DNA sequence assembly program**. *Genome Res* 1999, **9**:868–877.
68. Ruepp A, Zollner A, Maier D, Albermann K, Hani J, Mokrejs M, Tetko I, Guldener U, Mannhaupt G, Munsterkotter M, Mewes HW: **The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes**. *Nucleic Acids Res* 2004, **32**:5539–5545.
69. Cantón FR, Quail PH: **Both phyA and phyB mediate light-imposed repression of PHYA gene expression in *Arabidopsis***. *Plant Physiol* 1999, **121**:1207–1216.
70. Zhao S, Fernald RD: **Comprehensive algorithm for quantitative real-time polymerase chain reaction**. *J Comput Biol* 2005, **12**:1047–1064.
71. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F: **Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes**. *Genome Biol* 2002, **3**:R34.

doi:10.1186/1471-2229-12-100

Cite this article as: Villalobos et al.: Reprogramming of gene expression during compression wood formation in pine: Coordinated modulation of S-adenosylmethionine, lignin and lignan related genes. *BMC Plant Biology* 2012 **12**:100.

Submit your next manuscript to BioMed Central and take full advantage of:

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at
www.biomedcentral.com/submit



AGRADECIMIENTOS

En primer lugar quiero agradecer sinceramente a mi director, Francisco J. Ruiz Cantón, por haber estado a mi lado durante este largo camino. Es gracias a su asesoramiento, consejos, paciencia y palabras de ánimo que este trabajo ha podido realizarse. ¡Muchas gracias por haberme ofrecido esta gran oportunidad!

Agradezco también profundamente a mi director Francisco M. Cánovas Ramos por aceptarme en su grupo de investigación, por su ayuda con todo lo que necesitaba y su confianza en mí. Su punto de vista muchas veces me ha dado el empujón para continuar este camino apasionante. ¡Muchísimas gracias!

A Cañas quiero expresar mis sinceras gracias por abrir una ventana sobre el mundo de las asparraginasas. Tu trabajo previo me ha animado a continuar explorando este tema apasionante. ¡Gracias sobre todo por tus estimulantes críticas y apoyo incondicional!

Gracias también a Fernando por introducirme en el mundo de las proteínas, por su asistencia profesional con las técnicas y por todo lo que me ha enseñado. ¡Gracias por tu eterna alegría tan contagiosa!

A Belén por su apoyo tanto espiritual como técnico y por dejarme usar sus muestras de pinitos sometidos a condiciones de estrés.

A Vanesa por su gran amistad y su habilidosa ayuda con las preparativas para la espectrometría de masas.

A mis compañeros del grupo de investigación, Marina, Jorge, Maite, José Miguel y otros con quienes me he cruzado a lo largo del camino. Gracias por estar siempre dispuestos a ayudarme, aconsejarme y por vuestra amistad.

A Concha Ávila por su apoyo, consejos y comentarios constructivos.

A Reme por su generosa asistencia con todo lo que necesitaba.

Quiero agradecer especialmente a David Pacheco Villalobos por iniciarme en el trabajo de laboratorio hace tantos años. Su pasión por la investigación ha sido la semilla de mi propia pasión.

Gracias también a Juan Jesús por su eterna disponibilidad a ayudar a resolver mis dudas y problemas con gran profesionalidad. ¡Gracias por tu entusiasmo y apoyo!

A Nandi por ser una gran persona y una gran amiga. ¡Gracias por tu interés, tus consejos y por tus palabras de ánimo!

A mis otros compañeros que han compartido el laboratorio conmigo a lo largo de estos años, Sara, Marco, Alicia, Saíd, Vicky y otros que han pasado por ahí, les agradezco por haberme siempre ayudado con todo lo que necesitaba, pero sobre todo por vuestra amistad. Habéis hecho que el laboratorio ha siempre sido un lugar agradable al cual da ganas de venir a trabajar.

Como último, pero no menos importante, agradezco a mi amor Erwin, por compartir conmigo su entusiasmo para la ciencia, por su perseverancia contagiosa e apoyo infinito. ¡Te quiero mucho!